

University of Groningen

## Luxatie-fracturen van de enkel, opereren of niet opereren : methodologische aspecten van de beslissingsprocedure

Visser, Gabriel Jan Petrus

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*  
1975

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Visser, G. J. P. (1975). *Luxatie-fracturen van de enkel, opereren of niet opereren : methodologische aspecten van de beslissingsprocedure*. [, Rijksuniversiteit Groningen]. [S.n.].

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

**GABRIEL J.P. VISSER**

# ***Luxatie-fracturen van de enkel***

**Opereren of niet opereren**

**Methodologische aspecten van  
de beslissingsprocedure**

## LUXATIE-FRACTUREN VAN DE ENKEL



## STELLINGEN

1

Het bepalen van het 'ventricular mixing volume' kan een belangrijke parameter zijn in de beoordeling van het operatierisico bij oudere patiënten.

2

Ieder integrerend ziekenhuisinformatiesysteem, dat niet begint met redefinitie van de kwalitatieve en kwantitatieve inhoud van de klinische gegevens, het opstellen van probleemformuleringen en het ontwerpen van nauwkeurige klinische onderzoekmodellen, zal, wat de wetenschappelijke bewerking van de klinische gegevens betreft, op een kostbare mislukking uitlopen.

3

Wetenschappelijke gegevensbewerking van volledige patiëntenpopulaties is noodzakelijk voor het verkrijgen van voortdurende beoordeling van de resultaten van het medisch handelen.

4

Voortdurende resultaatbeoordeling is noodzakelijk voor nauwkeurige besturing van het modern medisch handelen.

5

Nauwkeurig besturen van het modern medisch handelen is bovendien de natuurlijke rem op de kostenstijging van de gezondheidszorg.

6

Indien alleen artikelen met statistisch geldige conclusies tot publicatie werden toegelaten, zou de omvang van de medische literatuur met ruim 75% afnemen.

7

Bij het bepalen van de waarde van een proefschrift kunnen de bijgevoegde stellingen geen enkele rol spelen.

8

Met behulp van de massaspectrograaf kan, ook bij patiënten die een hoge zuurstofconcentratie krijgen toegediend, een betrouwbare meting van de zuurstofconsumptie worden verricht.

9

Voor de opbouw van besturingssystemen voor het medisch handelen kan slechts van kleinere researchunits succes worden verwacht.

10

Bestuurlijk beleid, dat aandringt op collectieve samenwerking in de opbouw van besturingssystemen voor het medisch handelen, wijst op een naieve onbekendheid met de nog op te lossen problematiek.

11

Jaarlijks worden 14 autoinzittenden bij een ongeval maatschappelijk blind. Zij zouden dit letsel niet hebben opgelopen, indien de Nederlandse autovoorruiten van gelaagd glas waren voorzien.

12

Bij een strikt universitair gebeuren als een promotie zou de universitaire drukkerij tegen kostprijs voor de vermenigvuldiging van het proefschrift dienen zorg te dragen.

13

Er bestaat een discrepantie tussen de mate, waarin het Ministerie van Volksgezondheid en Milieuhygiëne het overnemen van patiënten uit het ziekenhuis door een verpleeghuis denkt te bevorderen en de mate waarop, volgens behoeftekriteria hiervoor, in het verpleeghuis bedden beschikbaar worden gesteld.

14

Zowel bij de diagnostiek als bij het instellen van een therapie schuilt vaak een adder onder het gras. Dit is vermoedelijk de reden dat het dier zo openlijk in ons vaksymbool staat opgesteld.

Stellingen behorende bij het proefschrift van G. J. P. Visser.

RIJKSUNIVERSITEIT GRONINGEN

**LUXATIE-FRACTUREN VAN DE ENKEL**  
**OPEREREN OF NIET OPEREREN**  
**METHODOLOGISCHE ASPECTEN VAN DE BESLISSINGSPROCEDURE**

PROEFSCHRIFT

TER VERKRIJGING VAN HET DOCTORAAT IN DE GENEESKUNDE  
AAN DE RIJKSUNIVERSITEIT TE GRONINGEN  
OP GEZAG VAN DE RECTOR MAGNIFICUS DR. M. J. JANSSEN  
IN HET OPENBAAR TE VERDEDIGEN OP WOENSDAG 8 OKTOBER 1975  
DES NAMIDDAGS TE 4 UUR

DOOR

**JAN PETRUS VISSER**  
geboren te Den Haag

PROMOTOREN : PROF. DR. P. J. KUIJER  
PROF. DR. W. MOLENAAR  
COREFERENT : DRS. B. BINNENDIJK



## EEN WOORD VOORAF

Dit proefschrift beschrijft een onderzoek naar de betekenis van de stabiele fixatie en functionele nabehandeling van luxatiefracturen van het bovenste spronggewricht. Het onderzoek werd verricht in het Gemeenteziekenhuis in Den Haag, destijds aan de Zuidwal.

De toegepaste operatietechniek werd nauwkeurig door Weber (1966) beschreven. Over de enkelfracturen bestaat een omvangrijke literatuur. De literatuuropgave van Golterman (1965) bevatte reeds ongeveer 600 titels, terwijl Jansen (1971) reeds uitvoerig aandacht besteedde aan de door Weber aanbevolen behandelingstechniek. In vele gevallen bevat de literatuur klinische studies betreffende de resultaten van de toegepaste behandeling. Meestal worden conclusies getrokken. Aan de bewijsvoering wordt vaak weinig aandacht besteed, de onderzoekgegevens worden op beschrijvend-statistische wijze besproken. Hierop volgt de conclusie al of niet na een statistische toetsing op de eindscores.

Freudenthal (1966) maakte de volgende opmerking: 'De wiskundige vraagt voor elke uitspraak een bewijs, niet omdat hij zijn collega's niet vertrouwt, maar omdat veelal het bewijs het enige middel is, om te weten te komen, wat met de uitspraak is bedoeld'. Mijns inziens is deze opmerking evenzeer van toepassing op uitspraken, die worden gedaan op basis van klinisch onderzoek.

Bij een poging tot exacte bewijsvoering op basis van de klinische gegevens kwamen een aantal bijna onoplosbare problemen naar voren. Deze waren vooral moeilijk oplosbaar omdat ze bij de opzet van het onderzoek niet waren voorzien. In dit verband wordt een naar verhouding grote aandacht besteed aan deels methodologische problemen. Ik meen echter dat daardoor de beoordeelbaarheid van de onderzoekconclusies aanzienlijk heeft gewonnen.

Aan al degenen, die mij gedurende de lange tijd, die het onderzoek in beslag nam, op de een of andere wijze ondersteunden, breng ik hier een woord van welgemeende dank.

Ik mag hierbij wel noemen:

Dr. J. D. Bom, mijn opleider in de chirurgie, die mij in de gelegenheid stelde alle daarvoor in aanmerking komende patiënten volgens de methode van Weber te behandelen. Hiervoor wil ik gaarne mijn erkentelijkheid betuigen.

Prof. Dr. P. J. Kuijjer, die bereid was als mijn promotor op te treden en die ik gaarne dank zeg voor zijn voortdurende belangstelling en daadwerkelijke hulp. Prof. Dr. W. Molenaar, die als mathematisch statisticus bereid was als copromotor op te treden. De talloze besprekingen, die wij hadden zijn van grote invloed geweest op de uiteindelijke vorm van het proefschrift.

Mijn collega J. K. Bakker, die veel van mijn opleiding voor zijn rekening nam en later zoveel belangstelling voor de toen nieuwe operatietechniek had, dat hij vaak op inconveniërende tijdstippen met de enkeloperaties meedeed.

De anaesthesisten K. Vonk en G. Kloppenburg. Voor hen betekende het invoeren van de operatie-indicaties van Weber een extra belasting van hun toch al zware taak. Zonder hun bereidwillige hulpvaardigheid was het onderzoek niet mogelijk geweest. De dames A. Vocking en H. Oostendorp-Wijns. Als secretaressen deden ze een grote hoeveelheid werk bij het uitzoeken van dossiers, oproepen van patiënten en controleren van adressen.

De heer G. Bal, die met grote nauwkeurigheid de voor het onderzoek belangrijke röntgenfoto's uitzocht en tegen zoekraken bewaakte.

Collega J. Bos wil ik gaarne bedanken voor zijn vertaling van de samenvatting.

De heer L. Th. van der Wee van het Universitair Rekencentrum Groningen, die als deskundig discussiepartner altijd bereid was de methodologische aspecten te bespreken, terwijl zonder het vooral door hem ontwikkelde WESP programma de daadwerkelijke gegevensbewerking zeer veel moeilijker was geweest.

Mijn collega's B. Binnendijk en H. Klasen wil ik van harte dank zeggen voor de wijze waarop zij mij in de gelegenheid stelden het proefschrift af te ronden.

## INHOUD

<b>Inleiding</b>	1
<b>Hoofdstuk 1 Bespreking van de behandelingsmethode van Weber</b>	5
A. Principes	5
B. Anatomische en functionele beschouwing	7
C. Indeling van de fracturen	13
D. Diagnostiek	19
E. Operatieindicaties	21
F. Voorbereiding tot operatie	22
G. Operatietechniek	23
H. De functionele nabehandeling	26
I. Het naonderzoek	27
<b>Hoofdstuk II Eigen onderzoek, directe resultaten en analyse hiervan</b>	29
A. Eigen onderzoek	29
B. De statistische beslissing	31
C. Onderzoekresultaten en analyse van de bevindingen	36
<b>Hoofdstuk III Foutenbronnen bij klinisch naonderzoek</b>	40
A. Onvergelijkbaarheidsintroductie door gebruik van ongelijke klassificaties	41
B. Verkeerd gebruik van hypothesen	41
C. Onduidelijke inhoud van de beoordelingsnorm	48
D. Fouten bij het kwantificeren van waarnemingen	52
E. Verwaarlozen van de factor tijd	55
F. Verwaarlozing van de invloed van selectie bij het vormen van klinische steekproeven	57
<b>Hoofdstuk IV De gegevens</b>	72

**Hoofdstuk V   Vergelijkbaarheid, representativiteit, de toetsingscriteria** . . 78

A.   De vergelijkbaarheid van de twee steekproeven . . . . . 78

B.   Zijn de te vergelijken steekproeven representatief? . . . . . 83

C.   De toetsingscriteria . . . . . 87

**Hoofdstuk VI   De betekenis van het tijdsverloop tussen fractuur en naonderzoek, de factor tijd** . . . . . 92

A.   De grafische voorstellingen . . . . . 93

B.   Overlappende tijdklasse . . . . . 96

C.   Correlatie tussen de tijd en de waarde van de naonderzoekvariabelen . 97

**Hoofdstuk VII   Bewerking en conclusies** . . . . . 99

A.   Inleiding . . . . . 99

B.   Vergelijking ten aanzien van de toetsingscriteria die het behandelingsresultaat samenvatten (de variabelen 97 en 175) . . . . . 101

C.   Nadere analyse van het gedrag van de somscores 175, 178 en 179 . 107

D.   Worden de weke delen door de operatieve behandeling belangrijk nadelig beïnvloed? . . . . . 111

E.   Is het mogelijk een optimale operatieindicatie uit de beschikbare gegevens te berekenen? . . . . . 113

F.   Samenvatting van de belangrijkste conclusies uit hoofdstuk VII . . 115

**Samenvatting** . . . . . 118

**Summary** . . . . . 125

**Literatuur** . . . . . 131

## INLEIDING

De prognose van een intraarticulaire fractuur is in het algemeen twijfelachtig. Sterker dan bij de extraarticulaire fractuur wordt het pijnloos functioneel herstel van de getroffen extremiteit bedreigd. Dit geldt niet in de laatste plaats voor de enkelfractuur. Hierdoor samen met zijn frequente voorkomen vormt de enkelfractuur een klinisch therapeutisch probleem, dat aanleiding geeft tot een grote hoeveelheid literatuur.

Een oud therapeutisch dilemma is de keus tussen een conservatieve en een operatieve behandelingstechniek. In 1966 verscheen een monografie over de letsels van het bovenste spronggewricht van de hand van Weber. Weber is één der leden 'van het eerste uur' van de inmiddels bekende, oorspronkelijk Zwitserse, 'Arbeitsgemeinschaft für Osteosynthesefragen' (A.O. groep). Door het werk van deze groep kwam de oefenstabiele osteosynthese als behandelingsmethode van fracturen sterk onder de aandacht. In zijn monografie brengt Weber dan ook de oefenstabiele osteosynthese bij enkelfracturen naar voren. Naast een uitvoerige beschrijving van de gebruikte operatietechnieken, geeft hij een eenvoudige en makkelijk te hanteren indeling van de fracturen van het bovenste spronggewricht.

Deze is vooral gericht op het bepalen van de prognose van het letsel en op het stellen van een operatieindicatie. Het belangrijkste is echter zijn betoog over de operatieindicaties zelf. Hij tracht deze te ondersteunen door hypothesen, die hij afleidt uit een aantal anatomische en functionele beschouwingen.

Zo gaat hij ervan uit, dat een gewrichtsletsel zo snel mogelijk moet worden geoefend teneinde een optimale genezingskans te krijgen. Daarom zou oefenstabiele fixatie van een intraarticulaire fractuur een absolute behandelingseis zijn.

De voorste vorkband (voorste syndesmose) speelt bij de fractuurindeling van Weber een centrale rol. Het effectief functioneren van deze structuur bepaalt, naar zijn inzichten, de prognose van het letsel. Is de voorste syndesmose niet meer intact, dan ontstaat er stabiliteitsverlies, hetgeen naast ongemak en klachten van de patiënt ook discongruentie van de kraakbeenvlakken ten gevolge heeft. Dit laatste zou aansprakelijk kunnen worden gesteld voor het optreden van posttraumatische arthrose. De verscheuring van de syndesmose kan op de röntgenfoto met grotere of kleinere zekerheid worden vastgesteld en hangt ten nauwste samen met de

hoogte van de fibulafractuur. Moet een verscheuring van de voorste syndesmose worden aangenomen dan is, zo stelt Weber, een absolute operatieindicatie aanwezig.

In zijn boek beschrijft Weber tweehonderd patiënten, die volgens zijn indicaties zijn geopereerd en die na gemiddeld 15 maanden zijn onderzocht. Hij vergelijkt de bij dit onderzoek gevonden resultaten, met literatuurgegevens betreffende conservatief behandelde patiënten en meent een flinke resultaat verbetering bij zijn patiënten te kunnen vaststellen.

Bij het lezen van de duidelijke aanwijzingen, die Weber in zijn boek geeft, komt men in de verleiding om zijn indicaties zonder meer over te nemen. Immers, iedere chirurg, die veel fracturen in zijn praktijk ziet, krijgt een aantal lelijke enkel-fracturen te behandelen, die een flinke invaliditeit achterlaten. Zo leert hij het therapeutisch probleem bij deze fracturen van nabij kennen. Het hanteren van de indicaties van Weber zou een aantal van deze dilemma's doen verdwijnen.

In het Gemeenteziekenhuis in Den Haag komen veel patiënten met enkelfracturen ter behandeling. Tot eind '66 was men gewoon deze in eerste instantie conservatief te behandelen. Pas bij het niet bereiken van een aanvaardbaar geachte stand langs conservatieve weg, werd tot operatie besloten. Meestal werd dan de mediale malleolus bloedig gereponeerd en gefixeerd. Om te bepalen of invoering van de indicaties en technieken van Weber tot een verbetering van het behandelingsresultaat zou leiden werd een onderzoek opgezet dat zowel uit een retrospectief als een prospectief onderdeel bestond. Het retrospectieve deel betreft patiënten, die in de jaren 1962, 1963 en 1964 in het Haagse Gemeenteziekenhuis waren behandeld wegens een enkelfractuur. Zij zouden voor een naonderzoek worden opgeroepen.

In het kader van het prospectieve onderdeel werden alle patiënten, die na januari 1967 ter behandeling van een enkelfractuur in dit ziekenhuis kwamen, volgens de principes van Weber behandeld. Ook deze patiënten werden later opgeroepen voor een naonderzoek.

Zo ontstonden twee series patiëntengegevens. De eerste was afkomstig van het retrospectieve deel van het onderzoek. De gegevens van deze serie zijn afkomstig van patiënten die volgens een Niet Operatieve techniek voor een enkelfractuur waren behandeld. Deze groep gegevens werd daarom de NO groep genoemd.

Aangezien de tweede serie gegevens werd verzameld bij patiënten die een Primair Operatieve behandeling ondergingen werd deze serie de PO groep genoemd. Om de techniek van Weber zoveel mogelijk recht te doen wedervaren werden alle patiënten uit de PO groep steeds door de auteur geopereerd. Variaties in operatietechniek werden op die manier zoveel mogelijk uitgesloten en onervarenheid met de nieuwe techniek kon alleen in het begin van het onderzoek een rol spelen.

Ook alle naonderzoeken werden door de auteur verricht.

Om de vergelijkbaarheid met de resultaten van Weber zo groot mogelijk te doen zijn werden alle fracturen naar zijn indeling geklassificeerd en werden alle gegevens, die

op het door Weber ontworpen naonderzoekformulier voorkomen, volgens zijn aanwijzingen vastgelegd.

Deze gegevens werden gebruikt om de resultaten van de NO groep met die van de PO groep te vergelijken. De resultaten hiervan zijn in hoofdstuk II beschreven.

Bij een nadere analyse van de naonderzoekmethode, die door Weber werd aan-gegeven, bleken hieraan nogal wat bezwaren te kleven. Vergelijkbaarheidsbeperkin-gen, problemen rond het tijdsinterval tussen fractuur en naonderzoek, onduidelijk-heid van de toetsingscriteria en selectieproblemen bleken het betrouwbaar beoor-delen van de behandelingsresultaten ernstig te belemmeren.

In hoofdstuk III ga ik op deze problemen nader in. Het zal blijken, dat nauw-keuriger normaliseren, strikter objectiveren en meer bewust hanteren van de mogelijkheden en beperkingen van statistische beslissingstechnieken een voorwaarde is om tot een meer betrouwbare beoordeling van de behandelingsresultaten te komen.

In de daarop volgende hoofdstukken heb ik getracht het beschikbare gegevens-materiaal naar deze inzichten te bewerken en de resultaten daarvan te belichten. In dit gedeelte wordt getracht, iedere bewering door een logisch statistische beslissing te laten ondersteunen.

Er zijn dan ook nogal wat uitkomsten van statistische toetsen ook in de tekst verwerkt. Ik ben mij ervan bewust, dat deze toetsingsuitkomsten niet iedere chirurgische lezer direct zullen aanspreken en heb dan ook steeds getracht de betekenis te formuleren, terwijl reeds in hoofdstuk II de algemene strekking van de overschrijdingskans  $P$  bij een statistische toets wordt besproken.

De literatuuropgave is bewust beperkt gehouden. De omvang van alleen al de literatuur van de letsels van het bovenste spronggewricht is zo omvangrijk, dat zelfs een summiere samenhangende weergave daarvan een omvangrijk boekwerk zou vergen.

De probleemstelling van het onderzoek is: 'Is de operatieve behandelingstechniek van Weber beter dan een conservatieve techniek of niet?'

Het onderzoek tracht op deze vraag een zo exact mogelijk antwoord te geven. In dit kader kunnen de meningen van anderen niet doorslaggevend zijn, aangezien alle in de literatuur te vinden meningen steeds weer een hypothetische basis hebben en relatief vrije interpretaties zijn van onderzoekbevindingen of algemene opvattingen. Weergave van een deel van de circulerende meningen zou altijd leiden tot een subjectieve keuze van de auteur, die de lezer niet op de juiste, objectieve, wijze voorlicht, zodat ik het aanhalen van anderen tot het minimum heb trachten te beperken.

In het streven naar een exact antwoord op de probleemstelling van het onderzoek leek het wenselijk de methodologische aspecten van de beoordelingsbeslissing te bespreken. De literatuur, die hierbij werd geraadpleegd en ter nadere orientatie kan dienen, werd samengavat in een apart tweede deel van de literatuurlijst.





## Hoofdstuk I

### BESPREKING VAN DE BEHANDELINGSMETHODE VAN WEBER

#### *A Principes*

In het gebied van de voet en de overgang van voet naar onderbeen ontstaan frequent letsels. Zowel wat de oorzaak als de prognose betreft is er een duidelijk verschil tussen de letsels van het overgangsgebied van onderbeen naar voet en de meer distale letsels.

De letsels van het overgangsgebied worden vrijwel altijd door indirect inwerkende krachten veroorzaakt. Bij gefixeerde voet beweegt het onderbeen met geweld, waardoor in het gebied van de spronggewrichten krachten ontstaan, die luxatie trachten te bewerkstelligen. Soms worden deze zo groot dat de belastbaarheidsgrenzen van de afremmende structuren worden overschreden en bandverscheuringen of fracturen optreden. De meer distale voet blijft bij dit soort beschadigingen meestal gespaard, doch is op zijn beurt kwetsbaarder ten opzichte van directe krachten, waarbij vaak een zwaar voorwerp op de voet terecht komt.

Hoe vaak beide typen letsels optreden wordt geïllustreerd door de volgende cijfers: In het Academisch ziekenhuis in Groningen werden van januari 1969 tot september 1971 820 fracturen van voet- en spronggewrichten behandeld. Fracturen van de tenen bleven hierbij buiten beschouwing. Van deze 820 fracturen betrof het 346 keer fracturen van metatarsus of tarsus. Al deze fracturen waren door direct inwerkende krachten ontstaan. De 474 andere fracturen waren letsels van het bovenste spronggewricht, die in 459 gevallen door indirect geweld waren ontstaan. Slechts 15 letsels van het bovenste spronggewricht waren door direct inwerkende krachten veroorzaakt.

Ook wat de prognose betreft onderscheiden de letsels van het bovenste spronggewricht zich van de meer distale letsels.

Bij de laatsten hebben vooral de tarsusfracturen een matige prognose. Bij de bovengenoemde 346 distale fracturen waren 52 tarsusfracturen. Ook de distale fracturen, die met (sub)luxatie gepaard gaan, hebben een matige prognose, doch zij komen nog zeldzamer voor.

De letsels van het bovenste spronggewricht hebben, globaal gesproken 30% kans op onvolledig herstel (Borchardt 1914, Hansen 1919, Magnussen 1944, Reimers 1953, Dinstl 1963).

Dit alles illustreert het feit, dat de verbinding tussen talus enerzijds en tibia/fibula complex anderzijds klinisch een speciaal probleem vormt.

Immers letsels in dit gebied komen vaak voor, ze hebben een bijzondere indirecte genese en laten frequent restinvaliditeit achter.

Mechanisch is het gedrag van dit gebied gemakkelijk te begrijpen. Het betreft een verbinding tussen twee loodrecht op elkaar staande structuren: enerzijds het verticaal opgerichte onderbeen en anderzijds de parallel aan de bodem verlopende voet. Ter plaatse van de verbinding is een ingewikkelde scharniermogelijkheid. Iedere dergelijke constructie is kwetsbaar ten aanzien van krachten, die op de benen van het scharnier aangrijpen.

In de door Weber geadviseerde behandelingsprincipes voor de letsels van het bovenste spronggewricht zijn de door de Zwitserse Arbeitsgemeinschaft für Osteosynthese Fragen gepropageerde grondbeginselen voor fractuurbehandeling duidelijk terug te vinden. Deze zijn:

- Nauwkeurige repositie
- Stabiele fixatie, gericht op het bereiken van een primaire ossale doorbouw van de fractuur en
- Zo spoedig mogelijk beginnen van een actief oefenprogramma van de getroffen extremititeit.

De enkelfracturen dienen te worden beschouwd als gecombineerde letsels van ligamenten, kapsel en fracturen ter hoogte van het bovenste spronggewricht. Deze combinaties komen in een grote variatie voor. Op welke wijze fracturen, ligament- en kapsellaesies gezamenlijk zullen optreden hangt af van de stand van de voet tijdens het ontstaan van het letsel en van de richting van het inwerkend geweld. Er zijn een aantal steeds weer optredende combinaties te herkennen, waarbij een bepaald type fractuur samengaat met een vaste combinatie van ligamentlaesies en andere fracturen. Op basis van deze herkenbare combinaties van beschadigingen van het bovenste spronggewricht zijn talloze indelingen van deze letsels opgesteld. Iedere indeling heeft ten doel om een aantal typen enkelfracturen te kunnen definiëren. Het lijkt plausibel, dat niet ieder type dezelfde prognose heeft en tevens dat de optimale behandeling voor de diverse typen verschillend kan zijn. Daarom wordt getracht door middel van de indelingen de behandelingen beter te individualiseren en de operatieindicaties te specificeren.

Weber geeft een indeling van de enkelfracturen, die vooral gericht is op de aanwezigheid van een ruptuur of andere insufficiëntie van de voorste vorkband aangezien hij meent dat het intact zijn van deze verbinding tussen tibia en fibula

voor een groot deel de prognose van het letsel bepaalt. Deze hypothese leidt hij af uit een aantal anatomische en functionele beschouwingen. Een bespreking hiervan mag dan ook niet achterwege blijven.

## **B Anatomische en functionele beschouwing**

De anatomische verhoudingen van de verbindingen tussen voet en onderbeen zijn bijzonder gecompliceerd. Functioneel gezien vormen zij een eenheid en is het theoretiseren over slechts één onderdeel ervan kunstmatig.

Indien de interdigitale gewrichten buiten beschouwing worden gelaten in de veronderstelling, dat ze slechts een ondergeschikte rol spelen, dan kunnen aan de voet 30 verschillende gewrichten worden onderscheiden. Hoewel aan een aantal van de voetgewrichten een uitgesproken kromming van de gewrichtsvlakken is te ontdekken, zoals b.v. bij alle gewrichten van de eerste twee stralen van de tarsus en metatarsus, valt het bij de meeste van de andere voetgewrichten op, dat de ten opzichte van elkaar beweeglijke gewrichtsvlakken nagenoeg vlak zijn. Een duidelijke draaiingsas kan er, althans wat het benig skelet betreft niet aan worden herkend. De bewegingsmogelijkheid van deze gewrichten wordt blijkbaar in hoofdzaak bepaald door de vrijheid, die door de betrokken ligamentverbinding wordt toegestaan.

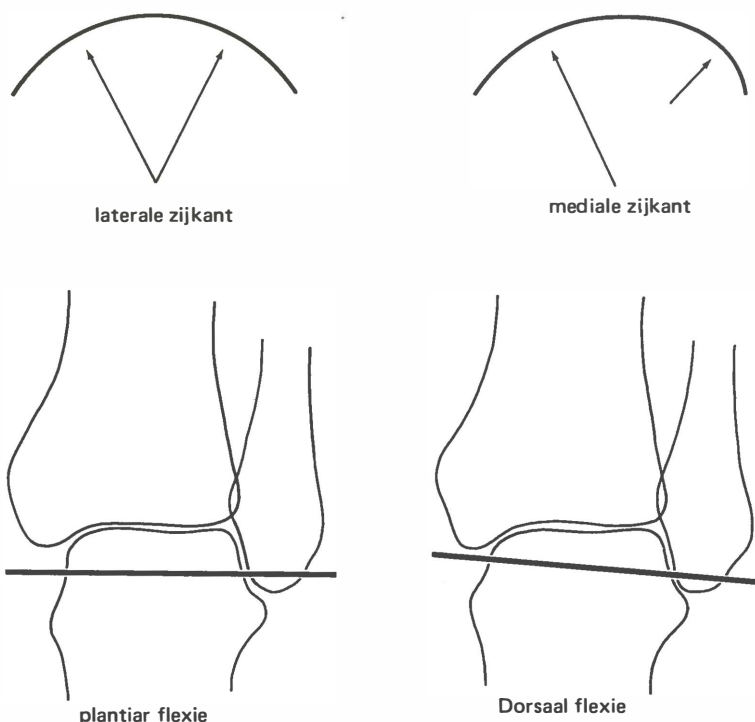
Een nauwkeurige bespreking van de anatomische verhoudingen van deze functionele eenheid valt buiten het bestek van een klinische studie en hoewel de eenheid van dit gebied daar geweld mee wordt aangedaan, zal de anatomische bespreking beperkt moeten blijven tot de verhoudingen, die de grondslag vormen voor de behandelings-techniek van Weber.

**B.1** Reeds uit de *benige vormelementen* valt af te leiden dat het bovenste spronggewricht geen eenvoudig scharniergewricht is. De gekromde zijkanten van de talusrol, waarbij mediaal een concave en lateraal een convexe kromming bestaat, zijn strijdig met een eenvoudige dorso-plantaire bewegingsmogelijkheid.

Men zou nog kunnen veronderstellen, dat de enigszins beweeglijke distale fibula zich aan het gekromde laterale taluszijvlak kan aanpassen, doch de mediale malleolus is star aan de tibia bevestigd en het gewrichtscontact aan de mediale zijkant zal dus door de beweeglijke talus moeten worden onderhouden.

Ook de kromming van het bovenvlak van de talusrol is minder eenvoudig dan het op het eerste gezicht lijkt. Aan de laterale zijrand ervan heeft de kromming ten naaste bij de vorm van een cirkelsegment.

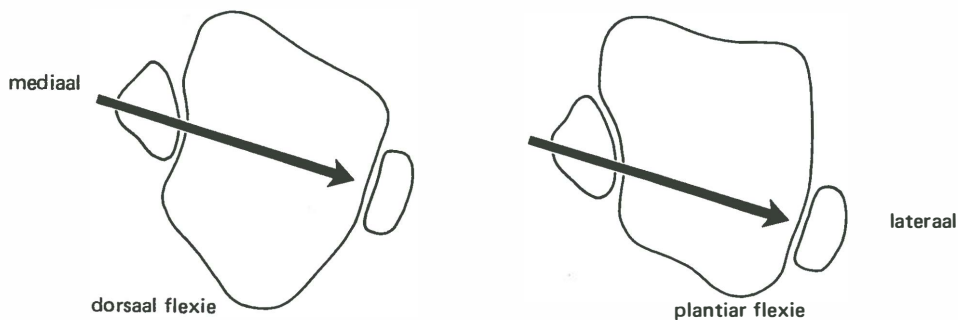
De kromtestraal is dus aan de voorkant gelijk aan die van de achterzijde van de kromming. Aan de mediale zijrand daarentegen is de kromtestraal aan de voorkant kleiner dan die van de achterkant, hierdoor zal de talus bij plantair-dorsaal flecteren een kanteling in de crurale vork moeten uitvoeren om goed contact met het distale



Figuur 1. Kromtestralen van de laterale en mediale zijrand van de talusrol (boven). Kantelbeweging van de talus in de enkelvork bij dorso-plantair flexie. (onder)

tibiaplateau te kunnen houden. Deze kanteling wordt uitgevoerd om een draaiingsas die een ventro-dorsale richting heeft en zich in de buurt van de laterale zijkant van de talus bevindt. Immers, zou hij zich ver van de laterale zijkant bevinden, dan zou ook deze zijrand ongelijke kromtestralen aan voor- en achterkant moeten vertonen. De reeds genoemde zijvlakken van de talus verdienen nog een nadere beschouwing. Zoals werd beschreven vertonen beiden een kromming, doch hierbij is tevens van belang dat:

- het effectieve gewrichtsvlak van de mediale zijkant slechts ongeveer even groot is als het corresponderende gewrichtsvlak van de mediale malleolus.
- het mediale gewrichtsvlak concaaf is gekromd.
- het mediale gewrichtsvlak meer ventraal is gelegen dan het midden van het laterale gewrichtsvlak.
- het effectieve gewrichtsvlak van de laterale zijkant ongeveer anderhalf keer zo groot is als het corresponderende gewrichtsvlak van de laterale malleolus
- het laterale gewrichtsvlak convex is gekromd en dat
- deze kromming niet gelijkmatig verloopt, doch ongeveer halverwege een richel vertoont.



Figuur 2. Bij plantair flexie glijdt de laterale malleolus over de richel in het laterale zijvlak van de talus, terwijl de talus een draaibeweging om de mediale malleolus maakt.

De genoemde krommingen van de zijkanten van de talus zijn alleen te begrijpen, indien de beweging van de mediale zijkant van de talus ten opzichte van de mediale malleolus veel kleiner is dan de verschuiving van de laterale taluszijkant ten opzichte van de laterale malleolus. Dit is alleen mogelijk indien de talus bij de plantairflexie beweging bovendien om de mediale malleolus draait. Dit zou niet alleen de concave kromming mediaal verklaren, doch tevens het verschil in effectieve gewrichtsovervlakken tussen de mediale en de laterale talus zijkant. Vastgesteld kan dan ook worden dat de talus bij plantairflexie een endorotatiebeweging in de crurale vork maakt (de Vogel 1970). Bij deze rotatiebeweging moet de laterale zijkant van de talus naar voren komen en als nu de fibulapunt door zijn ligamenten vrijwel niet ten opzichte van de talus van plaats kan veranderen, dan moet ook de fibula bij plantairflexie met de laterale zijkant van de talus enigszins mee naar voren bewegen. De talusrol is voor breder dan achter. Dit was van oudsher de reden, dat werd verondersteld dat de crurale vork bij dorsaalflexie breder zou moeten worden om de bredere talus voorzijde tussen de vorkpoten te doen passen. Door de endorotatie beweging om de mediale malleolus en door het feit dat de laterale malleolus verder naar achteren is geplaatst dan de mediale malleolus komt de brede voorkant van de talus in een heel ander licht te staan. De talus draait tussen de twee vorkpoten door en blijft gedurende de gehele beweging ongeveer even breed! (Zie fig. 2). Het is dus zeer de vraag of functioneel gezien het breedteverschil de fibula inderdaad noodzaakt om bij dorsaalflexie naar buiten te bewegen. Souquet (1965) weeserop, dat de uiterste standen van deze beweging nooit verder uit elkaar liggen dan één of twee millimeter, terwijl het breedteverschil van de talusrol vier tot zeven millimeter bedraagt. Dit doet veronderstellen, dat de fibula zich soms wel enigszins aan de breedte van het tussenliggende deel van de talusrol moet aanpassen, maar dat meestal door de onderlinge verhouding van de gewrichtsvlakken deze aanpassing niet nodig is.

Ook rotoire beweging van de fibula om zijn lengte as is vele malen vastgesteld. Weber meent, dat de fibula bij plantairflexie van de voet een exorotatoire beweging uitvoert; de meeste auteurs menen echter dat juist endorotatie plaats vindt.

Samenvattend voert de talus bij plantairflexie in de crurale vork tevens een kanteling uit in de zin van pronatie en een draaiing in de zin van endorotatie. Bij dorsaalflecteren treden de omgekeerde bewegingen op. De fibula moet bij plantairflexie naar boven en naar voren verschuiven ten opzichte van de tibia, terwijl hij tevens een rotatiebeweging uitvoert en soms enigszins naar binnen beweegt.

Belangrijk is het, om vast te stellen, dat deze bewegingen door de vorm van de samenstellende delen obligaat is gekoppeld aan de plantair/dorsaal flexie, zodat omgekeerd deze laatste gestoord zal worden bij blokkering van de fibulabewegingen b.v. door schroeffixatie van de fibula.

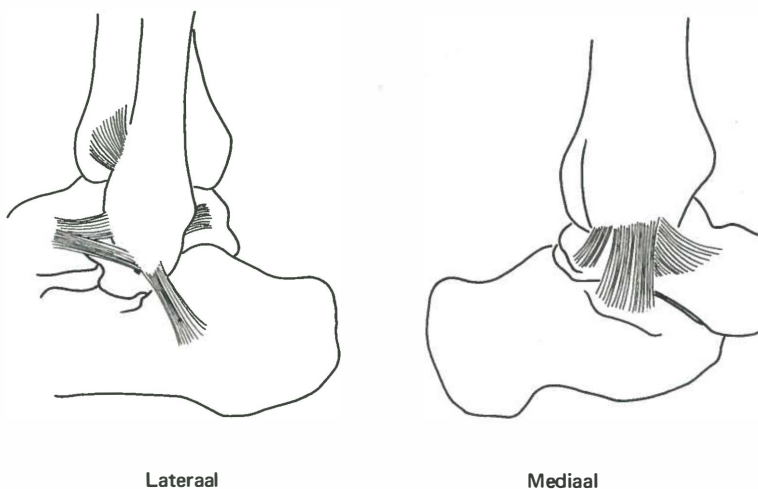
De fibula bewegingen worden vooral afgedwongen door de strafheid van de ligamenten. Het breedteverschil tussen voor en achterzijde van de talusrol speelt waarschijnlijk een minder belangrijke rol bij de fibula bewegingen dan meestal wordt verondersteld. De Vogel wees op de sturende functie van vooral de talo-fibulaire ligamenten. Eigen waarnemingen, verricht tijdens operaties konden geen uitsluitsel geven over de kleine excursies van de fibula, noch wat betreft het naar buiten bewegen, noch wat betreft de door Weber beschreven rotaties. De omstandigheden tijdens een operatie lieten geen fijne metingen toe zonder de steriele omstandigheden in gevaar te brengen, terwijl de omgevende weke delen verhinderden een goed inzicht te verkrijgen in de werkelijk door de benige gewrichtsdelen uitgevoerde bewegingen.

**B.2 De ligamenteuse structuren** (fig.3) aan beide zijkanten van het bovenste spronggewricht overspannen zowel dit bovenste als het onderste spronggewricht, terwijl het mediale ligamentum deltoideum bovendien aan het os naviculare insereert. Dit suggereert nog eens dat functioneel gezien het bovenste spronggewricht geen aparte eenheid is.

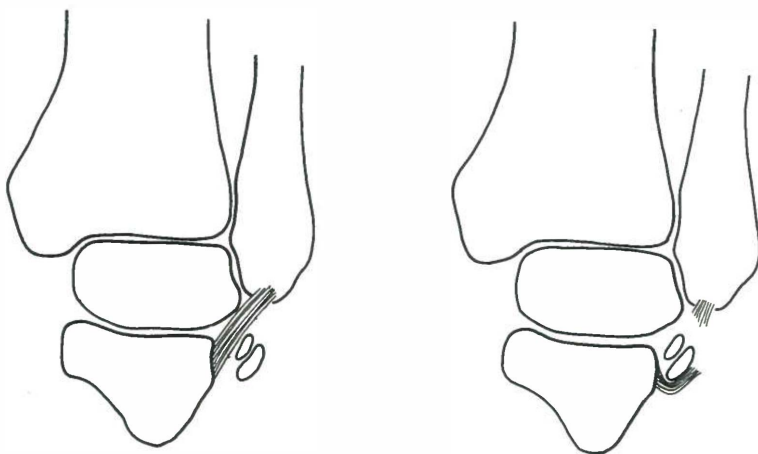
De vier samenstellende delen van het mediale ligament zijn in twee etages verdeeld, waarvan de aan de talus insererende delen dieper liggen dan de aan calcaneus en naviculare aangehechte delen.

Het laterale bandapparaat bestaat uit drie onderling apart ontwikkelde structuren. Delen van de voorste en achterste fibulo-talaire verbindingen zijn gedurende de gehele beweging van de talus aangespannen. (de Vogel 1970), ze fixeren de punt van de fibula en sturen hem in zijn bewegingen. Klinisch is het van belang, dat het lig. calcaneofibulare in tegenstelling tot de andere twee laterale ligamenten geheel apart van het gewrichtskapsel loopt, zeer stevig is en bij ruptuur vaak van zijn oorspronkelijke diepere ligging oppervlakkig van de peroneus pezen komt te liggen. Drioz (1956) wees er dan ook op, dat zonder het ligament in zijn oorspronkelijke positie terug te brengen, hechten niet mogelijk is. Eigen waarnemingen bevestigen dit. (fig. 4).

De tibio-fibulaire verbindingen worden gevormd door de voorste en achterste vorkbanden (syndesmoses), alsmede door het sterke lig. interossea dat naar boven over-



Figuur 3. De ligamentverbindingen ter hoogte van het BSG. De voorste syndesmose is op het lateraal aanzicht duidelijk zichtbaar.



Figuur 4. Normaal vertoont het lig. calcaneo-fibulare een diepe ligging ten opzichte van de peroneuspezen (links), na ruptuur komt het distale deel van het ligament soms oppervlakkig van de pezen te liggen (rechts).

gaat in de membrana interossea. Deze is 7 à 8 cm boven het gewricht weer versterkt (Souquet).

De achterste vorkband bestaat uit een ligamentplaat, die breed op de achterzijde van de tibia aanhecht. De onderste transversaal verlopende vezels zijn met kraakbeen bezet en nemen aan de gewrichtsvorming deel. Dit deel wordt nogal eens als een apart ligament onderscheiden.

De voorste syndesmose is veel kleiner en individueel verschillend sterk ontwikkeld. De vezels lopen vanaf de fibula schuin naar boven om te insereren aan de laterale onderkant van de voorzijde van de tibia.

**B 3. Belasting van het bovenste spronggewricht.** Weber stelt op grond van de uitkomsten van de onderzoeken van Braune en Fischer (1895), later bevestigd en aangevuld door Eberhard, Enman et al. (1947) het laterale fibula-syndesmose complex centraal voor het opvangen van een zeer belangrijk deel van de belastingen in het gewricht bij het gaan. Het blijkt dat gedurende de standfases krachten in het bovenste spronggewricht optreden, die soms vele malen het lichaamsgewicht bedragen. Vooral de eerste standfase waarbij de voet de bodem treft, is kritisch. Weber vergelijkt deze fase met het landen van een vliegtuig. Volgens Fischer bereiken druk en schuifmomenten in deze fase maximale waarden. Ontstaan in deze fase door 'verzwikking' bovendien krachtige torsiekrachten dan is het complex laterale malleolus+syndesmoses+tibia-achterkant tegen de belasting niet meer opgewassen en ontstaan gemakkelijk malleolus-fracturen of bandrupturen.

Het is duidelijk, dat door de grote belasting van de beweeglijke fibula tijdens de loopfases de stabiliteit van de talusrol in de vork in gevaar kan komen indien de fibula onder invloed van deze belasting zou gaan uitwijken. Weber stelt dan ook, dat voor een goede functie en stabiliteit van het bovenste spronggewricht de vorkbanden van even groot belang zijn als de benige malleoli. Reeds Pott (1775) wees op de centrale plaats van de fibula in de functie van het bovenste spronggewricht.

Uit de anatomische en functionele beschouwingen haalt Weber de volgende punten naar voren:

- Het bovenste spronggewricht is gecompliceerd gebouwd. De talusrol wordt in de vork door malleoli en bandapparaat geleid. Beiden zijn van even grote betekenis.
- Het bovenste spronggewricht (B.S.G.) ondervindt niet alleen drukkrachten, maar ook scheer-, schuif-, rotatie- en valgiserende krachten.
- De grootte en richting van de belastende krachten hangt af van de bodemdruk. De belasting is groter dan veelal wordt aangenomen.
- Een aparte plaats neemt de laterale malleolus en de achterzijde van de tibia in. In de eerste helft der standfases worden beiden krachtig tegen de talus gedrukt, onderling verbonden door de syndesmoses. Deze komen dan ook onder aanzien-



lijke tractie te staan. De belastende drukkrachten vormen hier zeker een vijfde van de totale gewrichtsdruk (d.w.z. 20 tot 40 kg).

- Tijdens de eerste helft der standfase roteert de dwarse draaiingsas van het B.S.G. naar binnen, zonder het frontale vlak te bereiken. Met deze beweging is een valgenderende tendens verbonden die wordt afgeremd door het ligamentum deltoideum en door actieve spierspanning. Gelijktijdig werken de dorsaal-flexoren, de quadriceps femoris en iets later ook de lange voetspieren als schokbrekers, waarbij de schuif- en scheerkrachten door het fibula+syndesmoses+achterste tibiarend complex elastisch worden opgevangen. Tijdens het afstoten van de grond volgt rotatie naar buiten. Hierbij wordt de fibula ontlast en komen de grote axiale drukken direct op de draagvlakte van de distale tibia.
- Het onderste spronggewricht (O.S.G.) en de meer distaal gelegen voetgewrichten zijn voor de harmonische functie van grote betekenis. Ze mogen bij beschouwingen over de belasting van het B.S.G. niet worden vergeten.
- Het is van belang, bij deze opsomming toe te voegen dat bewegingen van de fibula ten opzichte van de tibia essentieel zijn voor de goede functie van het B.S.G.

Tenslotte dient er op deze plaats op gewezen te worden dat de zware belasting van het laterale complex consequenties kan hebben, die klinisch van betekenis zouden kunnen zijn. Immers de gewrichtsvlakken van het B.S.G. zijn ingewikkeld gevormd. Kleine verschuivingen van corresponderende gewrichtsvlakken ten opzichte van elkaar zullen dus een relatief grote vermindering van het contact oppervlak kunnen veroorzaken. Zouden er bovendien oneffenheden in de gewrichtsvlakken zijn, b.v. ten gevolge van niet volledig gereponeerde fracturen, dan wordt dit effect nog sterker. Het blijkt dat doorsnijding van de voorste syndesmose een lateralisatie van de talus ter grootte van enkele millimeters toestaat. Indien nu bij een letsel van het B.S.G. een insufficiëntie van de voorste syndesmose is ontstaan dan moet worden verwacht, dat dergelijke kleine verschuivingen van de talus, onder invloed van de grote laterale belasting tijdens de loopfases, inderdaad plaats vinden. Het is dan ook niet te verwonderen, dat posttraumatische arthrosis zo'n grote rol speelt bij de resttoestanden van een letsel van het B.S.G. Gezien de centrale rol die aan het laterale complex van het B.S.G. moet worden toegekend voor de functie van dit gewricht, dient vooral laterale arthrosis gezien te worden als een ernstige bedreiging voor een blijvend goede functie van het bovenste spronggewricht.

### **C Indelingen van de fracturen van het bovenste spronggewricht**

Grofweg kan men twee soorten fractuurindelingen onderscheiden:

- 1) Indelingen, die het ongevalsmechanisme trachten af te leiden uit de opgetreden combinatie van letsels. Deze indelingen benoemen de fracturen meestal naar het

ongevalsmechanisme dat verondersteld wordt verantwoordelijk te zijn geweest voor het letsel. Deze indelingen staan bekend als de 'genetische' indelingen.

2) Indelingen, die los staan van het ongevalsmechanisme doch alleen trachten de regelmatig optredende combinaties van letsels rond het bovenste spronggewricht zo nauwkeurig mogelijk te rubriceren. Hun doel is meestal om de behandelende arts zo nauwkeurig mogelijk te informeren over de aard en omvang van de aanwezige letsels van alle elementen van het gewricht.

**C.1** De *genetische indelingen* hebben ten doel door informatie te geven betreffende het ongevalsmechanisme behulpzaam te zijn bij het uitvoeren van de 'genetische repositie' waarbij door omkering van het ongevalsmechanisme getracht wordt een gesloten repositie te bewerkstelligen.

De ontwikkeling van deze indeling is reeds oud. Na Pott, die een enkelfractuur beschreef welke door abductie werd veroorzaakt, verschenen talloze publicaties waarin gewezen werd op de relatie tussen het inwerken van een bepaald geweld en het ontstaan van een specifieke enkelfractuur. Aanvankelijk werd vooral gezocht naar de richting, die de inwerkende krachten hadden om verklaringen te vinden voor het optreden van groepen van steeds weer ongeveer dezelfde combinaties van enkelletsels, totdat vooral Lauge Hansen en met hem de meeste Skandinavische auteurs op het belang van de stand van de voet op het moment van inwerking van het geweld wezen. Toch heeft ook nu nog iedere auteur zijn eigen ideeën over de toe te passen indeling. Het gevolg is dat bij weinig publicaties handelend over de enkelfractuur het klinische materiaal op precies dezelfde manier is ingedeeld.

Souquet schrijft hierover:.....l'étude comparative des différentes classifications proposées est délicate par l'emploi d'une terminologie variant suivant la nationalité ou la formation technique des auteurs....'

Vervolgens haast hij zich echter om weer een nieuwe klassificatie in te voeren! Het resultaat van deze verwarring is dat de diverse conclusies getrokken uit het bewerkte gegevensmateriaal nauwelijks met elkaar te vergelijken zijn.

Men zou zich kunnen afvragen of de thans bestaande toestand betreffende de indeling van enkelfracturen zijn doel niet voorbijschiet. Hoe interessant het verband tussen letsel en geweld ook is, de patiënt is slechts gebaat bij een zo juist mogelijke aanpassing van de behandeling op zijn individuele fractuur. De prognose van een bepaalde groep enkelfracturen en de invloed daarop van een bepaalde behandelings-techniek kan des te beter worden bepaald naarmate meer auteurs hun ervaring op dit gebied meedelen op een wijze die onderlinge vergelijking toelaat. Hierbij is normalisering van de indeling van de fracturen een eerste vereiste.

Een algemeen bruikbare indeling dient eenvoudig te zijn, zodat ook de niet diepgaand geïnteresseerde er gebruik van kan maken zonder dat hij iedere keer moet nakijken hoe het ook weer precies in elkaar zit, doch tevens dient de indeling zoveel

mogelijk conclusies toe te laten over de vermoedelijke uitgebreidheid van het letsel alsmede een suggestie in te houden over de meest passende behandelingstechniek.

De genetische indeling, van nut voor de gesloten repositietechniek, loopt vaak uit in ingewikkelde faseringen.

De richting van het inwerkend geweld kan nu eenmaal schier eindeloos gevarieerd voorkomen, terwijl ook de stand van de voet een groot aantal variaties kan vertonen. Het is dan ook niet te verwonderen, dat steeds weer fractures optreden, die met de gehanteerde indeling niet goed te begrijpen zijn, aangezien iedere bruikbare indeling een simplificatie van de werkelijkheid inhoudt. Golterman (1965) wees er dan ook terecht op, dat de voorgestelde indeling van Lauge Hansen onvolledig is. De alternatieve indeling van Golterman is echter zo gecompliceerd, dat deze onbruikbaar is voor algemene toepassing.

**C 2.** Aangezien *Weber* de conservatieve behandeling voor vrijwel alle typen enkel-fracturen afwijst, heeft hij ook geen behoefte aan een genetische indeling. Hij stelt wel een aantal criteria op waaraan een indeling moet voldoen:

- de indeling dient een klassificatie van de fractures aan de hand van het röntgen-beeld mogelijk te maken.
- aan de hand van de indeling dient het röntgenbeeld met één oogopslag informatie te geven over de patholoog-anatomische afwijkingen tot in alle details, waarbij het ons niet primair interesseert hoe het betrokken letsel is ontstaan.
- bij de indeling dient de pathologie van het fibula-syndesmose gebied in het centrum van de belangstelling te staan.

Zo komt *Weber* tot de volgende indeling van letsels van het bovenste sprong-gewricht:

#### I. LUXATIE FRACTUREN

- A. Enkelfractuur met fibulafractuur distaal van de syndesmose.
- B. Enkelfractuur met fibulafractuur op het niveau van de syndesmose.
- C. Enkelfractuur met fibulafractuur proximaal van de syndesmose.

#### II. 'STAUCHUNGS' FRACTUREN

- A. Communiteive distale tibiafractuur met fibulafractuur.
- B. Idem met talusfractuur.
- C. Communiteive distale tibiafractuur zonder fibulafractuur.

#### III. ANDERE ENKELFRACTUREN

- A. Onderbeenfracturen uitlopend in het bovenste spronggewricht.
- B. Kinderfracturen van het bovenste spronggewricht.
- C. Atypische fracturen.

Deze indeling is een modificatie van een indeling die door *Danis* (1948) is voorgesteld als een indeling op patholoog-anatomische grondslag.

### C.2.1. De luxatiefracturen

Verreweg de meeste fracturen van het bovenste spronggewricht kunnen bij deze eerste groep van de indeling van Weber worden ingedeeld. Het zijn allen fracturen, die veroorzaakt worden door indirect aangrijpende krachten, die luxatie in het gewricht trachten te bewerkstelligen. Indien de krachten in het B.S.G. groter worden dan de belastbaarheidsgrenzen van de afremmende structuren in dit gewricht dan ontstaat bandruptuur of fractuur, er kan dan worden toegegeven aan de luxerende kracht totdat weer een volgende remmende structuur wordt aangespannen.

Het feit dat lang niet altijd een (sub)luxatie op de ongevalsfoto kan worden waargenomen, neemt niet weg dat in al deze gevallen tijdelijke (sub)luxatie moet hebben bestaan, aangezien de letsels anders niet waren ontstaan. Het is dan ook terecht dat Weber al deze fracturen onder de naam luxatiefracturen samenbrengt.

Het indelen van deze fracturen in drie grondtypen is van eminent belang, stelt Weber, aangezien deze indeling direct samenhangt met het al of niet intact zijn van de voorste syndesmose en 'als een rode draad door de literatuur vervolgd kan worden, dat bij conservatieve behandeling de prognose van het type A fracturen het beste — van het type B fracturen minder goed — en van type C fracturen het slechtste is'. Dit is des te opvallender, stelt Weber, aangezien bij alle drie types alle mogelijke combinaties van letsels aan mediale, laterale en achterzijde kunnen voorkomen. Volgens hem moet de verklaring gezocht worden in de toestand van de syndesmose.

De samenhang tussen de indeling van Weber en de toestand van de voorste syndesmose is als volgt:

- A fracturen: de voorste syndesmose is intact
- B fracturen: de voorste syndesmose is soms intact, maar soms ook verscheurd.
- C fracturen: de voorste syndesmose is altijd verscheurd.

Het is overigens opvallend dat ook in de genetische indeling van Lauge Hansen de fibulafractuur in het centrum is geplaatst. Ook in deze indeling speelt de hoogte van deze fractuur een rol, terwijl de fracturen van het type A ook bij Lauge Hansen een aparte groep vormen als supinatie adductie fracturen. De andere groepen van de twee indelingen overlappen elkaar echter.

De A-, B- en C fracturen uit de groep luxatiefracturen van de indeling van Weber zijn niet alleen te herkennen aan hun fibulaletsel, doch hun karakteristieke fibulafractuur komt ook steeds in vaste combinatie met een aantal andere letsels van het B.S.G. voor.

Zo kan men op grond van het type fibulafractuur een verwachting uitspreken over de aard van de andere letsels.

Als de mogelijke letsels onderscheiden worden in een mediaal-, een lateraal-, een syndesmose- en een letsel aan de tibia achterzijde dan komen de volgende combinaties voor:

	TYPE A	TYPE B	TYPE C
SYNDESMOSE	intact	?	verscheurd
MEDIAAL	intact of ruptuur lig. delt. of vrij steil verlopende malleolusfractuur	intact of rupt. lig. delt. of fract. malleolus in niveau van het gewrichtsvlak	rupt. lig. delt. of fract. med. mall.
LATERAAL	ong. horizontaal verlopende fract. onder het gewrichts- vlak, of ruptuur ligament	schuin verlopende fractuur, ventro- distaal naar dorso proximaal. Ventraal beginnend in het gebied van de voorste syndesmose	fract. boven de voorste syndesmose of verscheuring van de proximale tibio- fibulaire verbindingen
TIBIA ACHTERKANT	intact of mediaal achterkant fractuur, vaak deel. uitmakend van de mediale malleolus	intact of kleiner of groter lateraal fragment op te vatten als een afscheuringsfractuur van de achterste syn- desmose.	kleiner of groter lateraal achterkant frag- ment, op te vatten als een afscheurings frac- tuur van de achterste syndesmose, soms zijn zowel de tuberc. ant. als de tuberc. post. en bloc van de rest van de tibia afgescheurd

Naast de beschreven mogelijkheden, komen allerlei variaties in de details voor. Zo kunnen syndesmoses insufficiëntie vertonen door ruptuur van het ligament, doch evenzeer kunnen afscheuringsfracturen optreden aan de aanhechtingsplaatsen. Hierdoor kan het operatief herstel van een syndesmoseletsel wisselen van een simpele hechting van een ligamentruptuur tot een soms moeilijke reïnsertie door middel van een hemiossile naad. Zo kan het herstel van de ligamentstructuren soms een lastige procedure zijn.

De fracturen kunnen meer of minder communiteef zijn. Impressie van een gedeelte van het distale tibiaplateau kan voorkomen.

Ondanks deze variaties blijft de indeling van Weber makkelijk toe te passen en geeft hij een duidelijk beeld van de te verwachten letsels. Er zijn slechts twee duidelijke bezwaren:

- In de indeling bestaat geen duidelijke plaats voor het geïsoleerde letsel van de mediale malleolus. Toch komt een dergelijke fractuur, die meestal halverwege de malleolus ongeveer dwars verloopt, regelmatig voor. Hij wordt veroorzaakt door een indirect aangrijpend letsel en hoort duidelijk thuis in de groep der luxatiefracturen.

- Ten aanzien van conservatieve behandeling verschaft de indeling onvoldoende informatie ter ondersteuning van de genetische repositie, of de te adviseren immobilisatiestand.

Weber stelt echter als eis dat bij het behandelen van een luxatiefractuur van het bovenste spronggewricht een eventueel syndesmoseletsel operatief verzorgd dient te worden. Kan dus conservatief een goede repositie worden bereikt, dan kan alleen bij een fractuur van het type A van operatie worden afgezien. In dit verband dient men zich te realiseren, dat deze fracturen volgens de genetische indeling tot de supinatie adductie letsels gerekend worden.

## C 2.2 De 'stauchungs' fracturen

Weber onderscheidt in deze groep eveneens drie types. Daarbij is wederom bij het C type de syndesmose altijd verscheurd. Bij dit type is de fibula intact. Het gehele complex voet+voetwortel+distale tibia is naar proximaal verschoven, waardoor bij intacte fibula zowel de syndesmose als de laterale ligamenten insufficiënt zijn. De typen A en B vertonen beide een fibulafractuur even boven de syndesmose, de syndesmose is op zich zelf intact en gehecht aan een fragment van de tibia in het gebied van het laterale deel van het distale plateau. Hierbij is bij het type A de talus intact terwijl bij het type B de talus een fractuur vertoont. Het lig. deltoideum is nimmer insufficiënt, de fibulaire ligamenten alleen bij de fracturen van het C type.

## C 2.3 De andere enkelfracturen

### 1) Onderbeenfracturen uitlopend in het B.S.G.

Hierbij onderscheidt Weber:

Type A. geïsoleerde tibiafracturen, waarbij een syndesmoseletsel zeer wel mogelijk is en

Type B. fracturen van tibia en fibula, waarbij een syndesmoseletsel onwaarschijnlijk is.

### 2) Enkelfracturen bij kinderen.

De nadere indeling van deze fracturen houdt zich in belangrijke trekken aan die, welke door Aitken (1936) werd voorgesteld.

### 3) Atypische enkelfracturen.

Onder deze noemer voegt Weber een aantal letsels op het niveau van de meer distale voetwortel en middenvoet samen.

**C 3. Samenvatting.** Verreweg het grootste deel der enkelfracturen, die in een chirurgische kliniek een probleem kunnen vormen, laten zich onderbrengen in de eerste groep van de Weber indeling onder de naam luxatiefracturen. Deze indeling is gemakkelijk te onthouden door de simpele wijze, waarop de hoogte van de fibulafractuur als enig criterium wordt gebruikt.

Toch geeft deze indeling, althans met het oog op operatieve therapie, een aantal waardevolle aanwijzingen over de aard van de te verwachten laesies. Hiervan zijn de voornaamste:

- Bij de luxatiefracturen is bij de A fracturen de syndesmose intact, bij de C fracturen is hij zeker insufficiënt, terwijl hij bij de B fracturen soms intact is, maar ook insufficiënt kan zijn.
- Bij het röntgenologisch geïsoleerde laterale letsel, dient een ruptuur van het mediale lig. deltoideum te worden uitgesloten.
- Bij een schijnbaar geïsoleerd letsel aan mediale zijde, dient een letsel van het ligamentum fibulare, alsmede een hoge fibula fractuur te worden uitgesloten. In het laatste geval heeft men te maken met een type C letsel en moet ruptuur van de voorste syndesmose worden aangenomen.
- Tibiafragmenten aan de achterzijde bevinden zich bij de A fracturen aan de mediale zijde, terwijl ze bij de B en C fracturen op te vatten zijn als afscheuringsfracturen van de achterste syndesmose en zich dienovereenkomstig aan de laterale achterzijde van de distale tibia bevinden.

Voor de duidelijkheid zij nog eens opgemerkt dat deze aanwijzingen steeds betrekking hebben op de groep luxatiefracturen.

#### **D Diagnostiek**

1) De fracturen worden gediagnostiseerd door middel van de gebruikelijke twee standaard röntgenopnamen van het enkelgewricht. Om goede projecties te krijgen, dient er voor te worden gezorgd dat de voet bij de voor-achterwaartse opname ongeveer 20° naar binnen wordt gedraaid, omdat anders de benige gewrichtsdelen over de gewrichtsspleet heen geprojecteerd worden.

2) De rupturen van de mediale en laterale ligamenten worden gediagnostiseerd met behulp van 'kantelfoto's', waarbij wordt getracht de talus met de hand uit de enkelvork te kantelen, terwijl de tibia met de andere hand wordt gefixeerd. De hoek die de bovenzijde van de talusrol maakt met het distale tibia plateau is een maat voor de ernst van de ligamentverscheuring. Hoewel deze hoek uiteraard ook afhangt van de spanning van de remmende musculatuur, meent Dziob (1956) dat bij gebruik van locale anaesthesie, waarvoor hij 1% procaïne gebruikt, de bevindingen voldoende betrouwbaar zijn. Ook Weber maakt deze foto's met behulp van locale

anaesthesie (5cc procaïne 1%). Hij stelt bovendien, dat de behandelende arts het uitkantelen van de talus zelf moet verrichten, aangezien slechts hij precies weet waar het om gaat en röntgenfoto's, waarbij het uitkantelen van de talus door het röntgenpersoneel is bewerkstelligd, naar zijn ervaring onbetrouwbare resultaten geven.

Aangezien het infiltreren van een eventueel operatiegebied met een lokaal anaestheticum een minder gewenste techniek is en bovendien de onbetrouwbaarheid van de foto's samenhangt met een 'défense mechanisme' van de spieren, werden foto's gemaakt zonder anaestheticum, uitgaande van de gedachte, dat bij voorzichtige, lang aangehouden, tractie aan deze spieren de kanteling toch optreedt, aangezien het 'défense mechanisme' het opgeeft of niet optreedt.

In de gevallen waar later werd geopereerd, kregen we de gelegenheid om de verkregen kanteling te vergelijken met die, welke onder narcose mogelijk was (beeldversterker). Het bleek daarbij dat de primair gemaakte kantelfoto's een betrouwbaar beeld hadden gegeven. Indien de standaard opnamen een afwijking aantonen die een operatieindicatie inhoudt dan worden de kantelfoto's niet gemaakt, aangezien het controleren van de ligamenten dan met behulp van de beeldversterker onder narcose kan plaats vinden, zodat de toch wel onaangename ingreep achterwege kan blijven. De bestralings belasting van de chirurg behoeft bij deze methode niet groter te zijn dan bij het maken van kantelfoto's.

De interpretatie van de kantelfoto's is niet altijd even gemakkelijk. Vooral bij kanteling tot  $15^{\circ}$  is het moeilijk om met zekerheid vast te stellen of er een verse bandverscheuring aanwezig is. De locale klinische verschijnselen geven dan belangrijke informatie.

Een kapselscheur zonder echte instabiliteit kan echter ook veel locale verschijnselen in de vorm van pijn en haematoom geven, terwijl een duidelijke bandverscheuring soms met verrassend weinig locale verschijnselen gepaard gaat. Bonnin (1944) stelt, dat zonder verscheuring aan laterale zijde soms een hoek van  $5-15^{\circ}$  kan ontstaan. Berust een dergelijke instabiliteit op individueel 'slappe' banden, dan zal dezelfde instabiliteit ook aan de niet getroffen zijde moeten bestaan. Steeds moeten kantelfoto's dan ook van beide enkels tegelijkertijd worden gemaakt, zodat altijd vergelijken van de getroffen met de gezonde zijde mogelijk is.

Dziob meent dat bij kanteling aan de laterale zijde van:

- 5— $15^{\circ}$  een verscheuring van de voorste talo-fibulaire band kan bestaan.
- 15— $30^{\circ}$  waarschijnlijk een verscheuring van ook het lig. calcaneo-fibulare bestaat terwijl
- $>30^{\circ}$  de talus alleen kan worden uitgedraaid indien ook de achterste talo-fibulaire band insufficiënt is.

Voor kantelingen aan de mediale zijde kan een dergelijke scherpe indeling niet worden gemaakt. Vergelijkende kantelfoto's, samen met de klinische bevindingen



moeten tot een conclusie leiden over het al of niet intact zijn van het ligament.

3) Indien uit de standaard opnamen alleen een mediaal letsel kan worden gediagnostiseerd, terwijl aan laterale zijde geen bandletsel lijkt te bestaan, mag niet worden vergeten een nauwgezette klinische en röntgencontrole te verrichten van het meer proximale gedeelte van het onderbeen. In een aantal gevallen zal er dan een proximale fibula fractuur worden gevonden, die er op wijst dat een fractuur van het type C bestaat, welke gepaard gaat met een syndesmoseletsel.

4) Indien er verdenking bestaat op een letsel in het gebied van het onderste spronggewricht, dan kunnen halfzijdigse scheve beelden, die inzicht geven in de sinus tarsi, samen met een dorso-plantaire opname van de middenvoet, inlichtingen geven over botafwijkingen in dit gebied.

5) De syndesmoseletsels worden gediagnostiseerd, door ze via de indeling van Weber af te leiden uit de andere letsels, die zich betrouwbaar door de röntgenfoto's laten vaststellen.

Tallose technieken worden in de literatuur aanbevolen voor directe diagnostiek van een syndesmoseletsel. Bij het maken van een aantal controle opnames van gezonde enkels volgens de aanbevolen technieken, bleek steeds dat ook bij gezonde enkels resultaten kunnen ontstaan, die volgens de opgestelde criteria op een syndesmoseletsel wijzen. Alleen arthrografie vormt een uitzondering, hoewel ook hierbij een scheur in de voorste kapsel moeilijk van een syndesmoseletsel te onderscheiden is. Het geïsoleerde syndesmoseletsel, dat niet uit andere letsels kan worden afgeleid blijft zo een moeilijk met zekerheid te stellen diagnose. Operatief herstel van een dergelijke afwijking lijkt weliswaar wenselijk, doch het is slechts zelden mogelijk de indicatie te stellen.

## **E De operatie-indicaties**

Volgens Weber is iedere beschadiging van het B.S.G. die tot één of meerdere bandletsels en/of fracturen heeft geleid een absolute indicatie tot operatieve verzorging. Op iedere leeftijd van de patiënt is het volgens hem noodzakelijk om anatomische reconstructie te verkrijgen, zodat zowel bij het kind als bij de jongere volwassenen of oudere patiënt operatief moet worden gereponeerd.

Als contra-indicaties noemt hij de volgende punten:

- a. Contra-indicaties, die geen relatie hebben met het ongeval.
  - 1) voortkomende uit de algemene toestand, zoals decompensatio cordis, onregelde diabetes e.d.
  - 2) voortkomende uit de locale toestand, zoals ulcus cruris, slechte locale arteriële doorbloeding.
- b. Contra-indicaties, voortkomende uit het ongeval.
  - 1) algemene toestand, zoals shock

- 2) multipele letsels. Hier is steeds van geval tot geval te individualiseren of de malleolusfractuur de prioriteit moet hebben en of hij zo mogelijk in dezelfde narcose na of tijdens de verzorging van andere letsels geopereerd kan worden.
- 3) locale toestand. Als de locale toestand ongunstig is, dan wordt de operatie enige dagen uitgesteld. Indien reeds beschadiging van de weke delen aanwezig is, zoals ischaemie, contusie of blaarvorming, dan wordt de voet hoog neergelegd en afgewacht. Er wordt dan echter geopereerd zodra dat mogelijk is. Indien alleen de mediale zijde letsel vertoont van de weke delen, dan wordt lateraal wel geopereerd. In het algemeen kan men binnen de acht uren grens zonder extra risico direct opereren.

Gecompleceerde fracturen worden door Weber zonder uitzondering onmiddellijk geopereerd.

Over het tijdstip van operatie stelt Weber dat dit moment nooit gunstiger is dan direct na het ongeval. Bot en weke delen zijn dan nog in een nauwelijks veranderde toestand en regressieve veranderingen zijn nog niet begonnen. Het verse enkelletsel is voor hem dan ook een spoedindicatie par excellence: dag of nacht, altijd wordt er onmiddellijk geopereerd.

Bovendien merkt hij op, dat het ook om andere dan zuiver technische redenen verstandig is om primair te opereren, aangezien iedere patiënt na enkele dagen ziekenhuisverblijf besmet is met resistente ziekenhuiskiemen, zodat bij spoedoperaties lastige wondinfecties alleen daarom al minder te vrezen zijn.

## **F Voorbereiding tot operatie**

Behalve het normale prae-operatieve onderzoek adviseert Weber:

- a. bij gecompleceerde fracturen mag de steriele, primair aangebrachte wondbedekking nooit worden verwijderd, voordat de patiënt in de operatiekamer is. Dit om de kans op besmetting met ziekenhuiskiemen zo klein mogelijk te houden.
- b. bij gesloten fracturen met een grote dislocatie moet deze zo snel mogelijk worden opgeheven, hetgeen neerkomt op het corrigeren van luxatiestanden. Hiermede wordt zo goed mogelijk het optreden van weke delen beschadiging door druk van binnen uit, voorkomen. Daarna kan verdere (röntgen) diagnostiek plaats vinden.
- c. bij de andere gesloten fracturen dient men na het eerste klinische onderzoek direct standaard röntgenfoto's te laten maken.
- d. De directe operatievoorbereiding vindt plaats in de voorbereidingskamer van het operatiecomplex.
- e. Pas daar wordt onthaard onder steriele omstandigheden.
- f. Het is niet toegestaan, de enkel of complicerende verwonding te inspecteren zonder mond-neus masker.

g. Na desinfectie (met desogen of merfenoplossing) wordt het operatiegebied afgedekt met plastic folie.

h. Een bloedleegteband (pneumatisch) is bij het binnentreden van de operatiekamer reeds aangelegd en wordt voor het neerleggen van het been opgepompt tot 450 mm Hg druk.

Tegen de letterlijke uitvoering van deze adviezen van Weber hadden wij op enkele punten bezwaren, zodat de volgende wijzigingen werden aangebracht.

a. Desinfectie wordt uitgevoerd met Jodium tinctuur 1%. Na het scheren wordt twee maal geïjodeerd, vervolgens afgedekt, waarna het operatiegebied nog eens wordt geïjodeerd. Als het operatiegebied goed droog is wordt het operatieplastic aangebracht.

b. Zo mogelijk werd steeds zonder bloedleegte geopereerd en wel om de volgende redenen:

- Bij goede techniek is tijdens een operatie à chaud het meest hinderlijke bloed te verwachten van het fractuurhaematoom, dat zich in alle hoeken en gaten heeft verzameld en bij iedere manipulatie weer naar boven komt wellen. Dit geldt vooral bij de meer uitgebreide letsels. Deze 'bloeding' is door het aanleggen van een bloedleegte niet te voorkomen.
- Het is denkbaar, dat een periode van circulatiestilstand in een vers geïjodeerd been, waarbij verondersteld moet worden dat een flink aantal vaatwanden zijn beschadigd, het ontstaan van thrombose in het diepe veneuse systeem in de hand werkt.
- Bij twee patiënten in onze geopereerde serie, trad na enkele dagen na de operatie ondanks hun jeugdige leeftijd (<30 jaar), een, gelukkig niet dodelijke, longembolie op.
- Sindsdien hebben wij van het gebruik van de bloedleegte afgezien, behalve bij het herstel van geïsoleerde bandletsels. Naar onze ervaring gaf dat technisch nauwelijks een vergroting van de problemen. Daarna is geen embolie meer opgetreden. Er dient echter te worden vermeld dat naast het weglaten van de bloedleegte tijdens de operatie, iedere patiënt direct postoperatief werd ingesteld op marcoumar.

## **G Operatietechniek**

Weber is in zijn boek uitvoerig ingegaan op ter beschikking staande technieken ter stabilisering van de diverse typen fracturen.

Ik kan daar slechts het volgende aan toevoegen:

**G 1. De botlaesies.** Het is typerend voor de enkelfracturen, dat ze in een dergelijke verscheidenheid voorkomen, dat het noodzakelijk is over alle technieken te beschikken. Vaak is het mogelijk om reeds aan de röntgenfoto te zien welke techniek vermoedelijk zal voldoen, men komt echter bij de enkelfracturen vaker dan bij andere fracturen voor verrassingen te staan. Men zal in die gevallen dus tijdens de operatie moeten kunnen beslissen en over de benodigde materialen de beschikking hebben.

De Zuggurtung is voor de communitieve fracturen van mediale of laterale malleolus de enige mogelijkheid om tot een stabiel resultaat te komen. Deze Zuggurtung heeft tot taak, de bij normaal functioneren door een bepaald deel van het bewegingsapparaat op te nemen tractie over te nemen op staaldraden. De betrokken structuur wordt dus wat de belasting betreft, als het ware kortgesloten.

Dit principe is door de Zwitserse A.O. (zie pag.1) naar voren gebracht als logische osteosynthese techniek bij structuren, die trekkrachten op moeten nemen. Vooral bij fracturen van patella en olecranon zijn operaties volgens dit principe goed uitvoerbaar. Ook osteosynthesemateriaal in de vorm van metalen platen kan volgens dit Zuggurtungsprincipe worden aangebracht.

Hoewel voor de wat proximaal gelegen fibulafacturen de Rushpen techniek goed kan worden toegepast, ontstaan vaak moeilijkheden doordat de fibulaschacht te nauw kan zijn om de pen toe te laten. Voorzichtig opboren van de schacht vanuit de fractuur is dan vaak beter dan fraisen met de vaak botte rushpriem.

Indien bij een syndesmoseletsel na het operatief herstel aan de stevigheid van de gehechte band wordt getwijfeld, adviseert Weber een fixatie d.m.v. een schroef, die tibia en fibula aan elkaar fixeert een eindweegs proximaal van het distale tibio-fibulaire gewricht. Hierdoor wordt de gehechte syndesmose tijdens de functionele nabehandeling niet in gevaar gebracht. De schroef dient zo te worden aangebracht dat de fibula niet stijf tegen de tibia is gefixeerd. Daartoe wordt de schroef nadat hij zover is aangedraaid dat de fibula juist klem zit op de tibia, ongeveer anderhalve slag teruggedraaid, zodat voldoende speling tussen tibia en fibula aanwezig blijft om de bewegingen tijdens het oefenen mogelijk te maken. Voordat krachtige bewegingen noodzakelijk zijn, (bij belasten van het been) dient de schroef te worden verwijderd. Belasten wordt als regel na ongeveer acht weken toegestaan, zodat voor die tijd de syndesmoseschroef onder lokaal anaesthesie wordt verwijderd.

Bij het vervolgen van de geopereerde patiënten bleek mij echter, dat in een aantal gevallen, enkele maanden na het verwijderen van de syndesmoseschroef een benige verbinding ging ontstaan tussen tibia en fibula. Na die waarneming hebben wij de syndesmoseschroef nog maar zeer zelden toegepast.

**G 2. De ligamentlaesies.** Conform de bevindingen van Weber vonden wij in een groot aantal der mediale bandrupturen een knoop en knoopsgat mechanisme, waarbij de mediale malleolus door de ruptuur subcutaan was komen te liggen en het lig. laciniatum zich in de mediale gewrichtsspleet bevond. Na repositie kan met geknoopte OO-nylon in het algemeen een stevige hechting verkregen worden op de door Weber beschreven wijze.

Het viel op dat in een aantal gevallen het ligament, onder de pees van de m. tibialis post. gelegen, was verscheurd, terwijl de oppervlakkig gelegen structuur intact was. Het is dus noodzakelijk dit dieper gelegen deel te inspecteren ook als de oppervlakkige etage van het ligament niet is verscheurd. Soms moet hiervoor dit oppervlakkige deel worden gekliefd, aangezien het anders niet mogelijk is de pees voldoende uit zijn bed te lichten om de diepere structuren à vue te krijgen.

Tenslotte is het enkelletsel met de ruptuur van het lig. deltoideum de uitzondering op de regel dat eerst het laterale complex dient te worden gereponeerd en gestabiliseerd. Het dikwijls interponerende lig. laciniatum verhindert n.l. een goede repositie aan laterale zijde, zolang de mediale interpositie niet is opgeheven.

In verband met de laterale banden verdient de techniek van Weber volgens welke een heen en weer verlopende naad wordt gebruikt, eveneens aanbeveling. Aandacht dient echter te worden besteed aan rupturen van de meer apart verlopende middelste verbinding tussen fibula en calcaneus, die nog al eens van positie is veranderd (zie pag. 10). In die gevallen dient hij apart te worden gehecht.

Syndesmose rupturen gaan zeer frequent gepaard met afscheuringsfracturen aan fibulaire zijde. Met dunne Kirschner draadjes zijn deze fragmenten goed op hun plaats te houden. De uitstekende delen van deze draadjes kunnen worden gebruikt om een eventuele Zuggurtungsdraad langs te leiden, waardoor de fragmenten stevig worden gefixeerd. De syndesmose aanhechting aan de tibia scheurt nog al eens glad langs de tibia af. In die gevallen dient een transossale naad te worden gebruikt. Deze is dikwijls zonder boorkanalen te verwezenlijken door het spongieuze karakter van het tuberculum tibiale anterius waardoor het mogelijk is met een kleine, gladde niet al te kromme naald direct door het bot te steken. Hoewel het aanbeveling verdient het verscheurde kapsel te hechten op grond van algemeen aangenomen behandelingsprincipes van gewrichtsletsels is er in een groot aantal gevallen een dergelijk weefselverlies van het kapsel dat het hechten ervan technisch niet goed mogelijk is. Ik heb niet vast kunnen stellen, dat open laten van het kapsel invloed heeft op het eindresultaat.

Volgens Weber komt een kleine afscheuringsfractuur van de achterste syndesmose na anatomische repositie van de andere gewrichtselementen voldoende op zijn plaats te liggen. Deze situatie wordt niet bedreigd door de functionele nabehandeling en behoeft geen verdere fixatie. Wanneer bij de afscheuring van de achterste

syndesmose een groter fragment van het tuberculum tibiale post. is betrokken, dan dient dat exact gereponeerd en gefixeerd te worden. De vraag hoe groot het stuk mag zijn, dat men zonder nadere verzorging mag laten liggen, heeft al vele malen aanleiding gegeven tot discussies. Weber neemt hierover een arbitraire beslissing en zegt, dat hoewel een goed criterium moeilijk is te geven, naar zijn inzicht en een los fragment van de achterzijde van de tibia, dat gelijk of groter is dan één vijfde van de sagittale tibiadoorsnede ter hoogte van het distale tibiaplateau, operatief dient te worden verzorgd. Dit geldt uiteraard eveneens voor de mediale achterkantfragmenten. De gebruikte techniek voor mediale en laterale fragmenten is verschillend. De mediale kunnen via de mediale incisie direct van achteren uit met een schroef worden gefixeerd, terwijl de laterale fragmenten met een juist gerichte schroef van de voorzijde van de tibia uit, eventueel via de laterale incisie, kunnen worden gestabiliseerd.

#### **H** *De functionele nabehandeling*

Vele auteurs hebben zich bezig gehouden met het effect van functionele nabehandeling, waarbij zo spoedig mogelijk na een extremitetletsel met een actief oefenprogramma van de getroffen extremitet wordt begonnen. De meesten menen dat deze behandeling aanzienlijke voordelen biedt ten opzichte van een langdurige (gips) immobilisatie.

De meeste meningen berusten echter op hypothesen, die afgeleid worden uit een aantal algemene beschouwingen dan wel uit resultaten op korte termijn. Voor zover er gegevens over vergelijkingen tussen de resultaten op langere termijn van beide nabehandelingsopvattingen beschikbaar zijn, lijkt het minder eenvoudig om een beslissing te nemen over de te adviseren behandeling.

Trueta toonde aan, dat de bloedvoorziening van een pijpbeen verbetert bij actieve functie van de omgevende spierkoker. De Zwitserse A.O. groep is een sterk voorstander van de functionele nabehandeling en stelt zelfs, dat dit in het algemeen beleid van fractuurbehandeling thuis hoort. De resultaten zijn dan ook indrukwekkend, doch alleen door de korte tijd die soms nodig is om, na een extremitetletsel weer een goed functieherstel te verkrijgen. Ook de A.O. verschaft echter geen gegevens over vergelijking op langere termijn.

Weber stelt, dat de zin van functionele nabehandeling van enkelfracturen ligt in het tot functie brengen van alle gewrichtskapsels van de gehele voet, terwijl tegelijk ook de onderlinge glijvlakken tussen de spieren en de pezen weer in gebruik worden gesteld, voordat verklevingen of schrompelingen hebben kunnen ontstaan.

Vroege mobilisering verbetert de circulatie terwijl het gewrichtskraakbeen zijn normale glijfunctie nodig heeft, waardoor het minder gemakkelijk tot regressieve veranderingen komt. Haematomen verdwijnen onder deze behandeling zeer snel; hierdoor wordt fibrose beperkt.

Volgens hem is de functionele nabehandeling het beste middel om posttraumatische dystrophie te voorkomen. Indien de reconstructie anatomisch is en de osteosynthese werkelijk stabiel, dan ondervindt de patiënt van het oefenen niet de geringste pijn, zodat ook hiervan geen activering van dystrophische processen te vrezen is. In de praktijk wordt de functionele nabehandeling als volgt uitgevoerd:

- Aan het slot van de operatie wordt de enkel door middel van een korte achterspalk geïmmobiliseerd.
- Reeds de eerste postoperatieve dag wordt deze spalk onder leiding van een physiotherapeut(e) afgenomen en de patiënt aangemoedigd de enkel in het B.S.G. te bewegen.
- Na vijf dagen wordt de spalk geheel verwijderd en krijgt de patiënt instructie om ieder uur vijf minuten met de enkel te oefenen. Hij wordt aangemoedigd om deze oefeningen met beide enkels tegelijk te doen waarbij hij moet trachten om met de geopereerde enkel dezelfde bewegingsexcursie te bereiken als met de gezonde enkel.
- Eveneens wordt na vijf dagen begonnen door middel van draaibewegingen de meer distale gewrichten in het oefenprogramma te betrekken. Ook deze bewegingen worden met beide enkels tegelijkertijd uitgevoerd.
- Er wordt naar gestreefd om ongeveer 10 tot 14 dagen postoperatief een compleet functieherstel te bereiken.
- Voor succes bij dit oefenprogramma is begeleiding door een physiotherapeut zeer belangrijk. Deze begeleiding dient minstens dagelijks doch bij voorkeur meermalen per dag plaats te vinden.

### I. Het naonderzoek

Het naonderzoekformulier, zoals dat door Weber werd opgesteld, spreekt voor zich. De puntenwaardering verdient nadere explicatie.

#### a. Pijn

Geen enkele last	= 0
Enige last bij sterke belasting	= 1
Enige last bij normaal lopen	= 2
Last bij beweging reeds zonder belasting	= 3
Spontane pijn in rust	= 4

#### b. Lopen

Alle kwaliteiten normaal (normale gang, hardlopen, hiel- en teengang, op binnen- of buitenrand lopen, hurklopen)	= 0
Eén gangkwaliteit verminderd	= 1
Vermindering van twee kwaliteiten met licht kreupelen	= 2
Duidelijk kreupelen	= 3
Sterk kreupelen, altijd stok nodig	= 4

c. Activiteit	
Volledige activiteit zowel binnen als buiten het beroep	= 0
Normaal binnen het beroep, daarbuiten echter enige beperking	= 1
Normaal binnen het beroep, echter opgeheven activiteit buiten het beroep	= 2
Verminderde activiteit binnen het beroep	= 3
Verandering van beroep nodig	= 4
d. Röntgenbeeld	
Anatomisch perfect zonder arthrose	= 0
Anatomisch perfect met enige verkalking van een ligament	= 1
Anatomische onregelmatigheid alleen mediaal	= 2
Anatomische onregelmatigheid lateraal, d.w.z. arthrose	= 3
Trapje aan de achterkant d.w.z. arthrose. Dystrophie	= 4
e. Bovenste spronggewricht	
Volledige functie links = rechts	= 0
Ongelijkheid van ten hoogste 10°	= 1
Ongelijkheid van meer dan 10°, doch dorsiflexie tot 90° mogelijk	= 2
Niet gefixeerde spitsvoet met dorsiflexie mogelijkheid tot 95°	= 3
In het B.S.G. vergaand verstijfde voet met hinderlijke spitsvoet	= 4
f. Onderste spronggewricht	
Volledige functie links = rechts	= 0
Ongelijkheid, die nauwelijks aantoonbaar is	= 1
Ongelijkheid niet meer dan de helft	= 2
Ongelijkheid meer dan de helft	= 3
Ankylose van het onderste spronggewricht	= 4

De scores in deze zes onderdelen worden opgeteld tot een totaal beoordelings somscore. Deze is de basis voor het uitspreken van het oordeel uitstekend, goed of slecht resultaat en wel op de volgende wijze:

Somscore = 0 = zeer goed

Somscore 1 of 2 = goed

Somscore 3 of hoger = slecht

Hoewel ik heb getracht het naonderzoekschema zo precies mogelijk te volgen, leek het al spoedig, dat het beoordelingsschema hier en daar aan duidelijke inhoud tekort schiet. Zo is b.v. niet duidelijk in hoeverre de score van punt c alleen afhankelijk is van het behandelingsresultaat van de enkelfractuur. Zeker zal bij deze score het lokaal geldende verzekeringssysteem en de aard van het beroep een rol spelen.



## Hoofdstuk II

### EIGEN ONDERZOEK, DE STATISTISCHE BESLISSING, RESULTATEN VAN HET ONDERZOEK EN ANALYSE HIERVAN

#### A *Eigen onderzoek*

Zoals reeds in de inleiding werd uiteengezet ontstond het eigen onderzoek uit de behoefte na te gaan in hoeverre de door Weber aangegeven behandelingsmethode van fracturen van het bovenste spronggewricht tot resultaatverbetering zou leiden vergeleken met de resultaten van het tot dusver gevolgde behandelingsbeleid.

Gegevens werden verzameld van twee groepen patiënten:

1) Van januari 1967 tot oktober 1968 werden alle patiënten die het Gemeente-ziekenhuis van Den Haag binnenkwamen, ter behandeling van een enkelfractuur, zo nauwkeurig mogelijk volgens de aanwijzingen van Weber behandeld. Ze werden geopereerd door steeds dezelfde operateur (de auteur) en na enkele jaren onderzocht. Zo werden 97 patiënten geopereerd, waarvan er 85 konden worden naonderzocht. De gegevens van deze 85 patiënten vormen samen een groep gegevens, die afkomstig is van patiënten, die Primair werden geOpereerd wegens een enkel-fractuur. Deze groep gegevens wordt daarom de PO groep genoemd.

2) Vóór januari 1967 werden de enkelfracturen gebruikelijk conservatief behandeld. Alleen indien langs conservatieve weg geen aanvaardbaar geachte repositie werd verkregen werd bloedig gereponeerd. Meestal werd dan volstaan met het reponeren en fixeren van de mediale malleolus. Gegevens werden verzameld van patiënten, die op deze wijze wegens een enkelfractuur waren behandeld in de jaren 1962, 1963 en 1964. Zo konden gegevens van 235 patiënten worden gevonden. Na januari 1967 werden deze patiënten opgeroepen voor naonderzoek. Er konden 156 patiënten voor een compleet naonderzoek worden bereikt. De gegevens van deze 156 patiënten vormen samen een groep gegevens, die afkomstig is van patiënten, die op Niet Operatieve wijze wegens een enkelfractuur waren behandeld. Deze groep gegevens werd daarom de NO groep genoemd.

De PO en de NO groep werden met elkaar vergeleken volgens de criteria van Weber.

		WEBER		EIGEN ONDERZOEK	
			LITT.GEG.	PO	NO
Totaal		% F		% F	% F
	Zeer goed	77 116		22 19	19 30
	Goed	9 14		42 36	20 31
	Slecht	13 20	40-50%	35 30	60 93
	Totaal	150		85	154
A-Fracturen					
	Zeer goed	82 23		38 3	22 5
	Goed	3 1		38 3	39 9
	Slecht	14 4		25 2	39 9
	Totaal	28		8	23
B-Fracturen					
	Zeer goed	78 40		18 7	14 10
	Goed	10 5		55 22	15 11
	Slecht	12 6		28 11	71 52
	Totaal	51		40	73
C-Fracturen					
	Zeer goed	75 53		17 5	10 2
	Goed	11 8		38 11	5 1
	Slecht	14 10	50-60%	45 13	86 18
	Totaal	71		29	21

Selectie criterium	Steekproeven				
	W	PO	NO	P	Mann Whitney
Totale steekproeven					
id.	X	X	X	P=0,001*	z= +2,99
A-fracturen				P<0,00003*	z= +7,53
id.	X	X	X	P=0,18	z= +0,91
B-fracturen				P=0,01*	z= +2,18
id.	X	X	X	P<0,0007*	z= +3,18
C-fracturen				P<0,00003*	z= +5,21
id.	X	X	X	P=0,004*	z= +2,64
A- versus B-fract.				P<0,00003*	z= +5,04
id.		X		P=0,21	z= +0,80
B- versus C-fract.				P=0,006*	z= +2,53
id.		X	X	P=0,14	z= +1,11
			X	P=0,11	z= +1,25

Tabel 1. (boven) Drie groepen patiëntengegevens geordend volgens de beoordelingscriteria van Weber: 1. De door Weber gepubliceerde groep (W) 2. De primair geopereerde groep (PO). 3. De niet geopereerde groep (NO)

Tabel 2. (onder) Toetsingsresultaten van vergelijkingen van de diverse groepen. De groepen die vergeleken werden zijn steeds met een X aangegeven

\* P< significantiegrens

Tenslotte werd de PO groep vergeleken met de door Weber gepubliceerde resultaten van zijn behandelingsmethode. De groep gegevens, die door Weber werden beschreven wordt aangeduid als W.

De resultaten van deze eerste vergelijkingen zijn samengevat in de tabellen 1 en 2. De beschrijvingen ervan zijn te vinden in het derde onderdeel van dit hoofdstuk.

## B De statistische beslissing

Volgens de criteria van Weber kan men de naonderzoekgegevens van een patiënt samenvatten in een eindscore. Deze zou volgens Weber een middel zijn om een gevonden resultaat zeer goed, goed of slecht te noemen.

In de twee rechter kolommen van tabel 1 zijn de gegevens uit de PO en NO groep, verdeeld volgens die eindscore van Weber, weergegeven. In de bovenste rij van genoemde twee kolommen staan de verdelingen van de totale PO en NO groepen vermeld. Ze zijn:

	PO	NO
Zeër goed	19	30
Goed	36	31
Slecht	30	93

Aangezien het de bedoeling van het onderzoek was om een uitspraak te kunnen doen over de voorkeursbehandeling van enkelfracturen naar aanleiding van de resultaten na operatieve dan wel na niet operatieve behandeling, terwijl bovenstaande cijfers de resultaten van de twee typen behandeling voorstellen, zullen deze cijfers de basis moeten vormen van de beslissing omtrent de gezochte voorkeursbehandeling. Op het eerste gezicht krijgt men de indruk dat er in de NO groep toch wel erg veel resultaten slecht moesten worden genoemd. Het vergelijken van de werkelijke aantallen is echter nogal eens moeilijk, vooral indien er een groot verschil bestaat tussen de totale aantallen van de te vergelijken groepen. Men moet dan steeds gaan terugrekenen of er naar verhouding veel of weinig van een der vergeleken groepen in een categorie goed of slecht zijn terecht gekomen. Daarom is het gebruikelijk geworden de verdelingen te presenteren nadat de verhoudingsgetallen zijn berekend steeds naar een (fictief) totaal aantal van 100. Doordat de te vergelijken groepen nu worden gepresenteerd in verdelingen, die zijn gebaseerd op onderling gelijke totale aantallen is de vergelijking op het eerste gezicht gemakkelijker.

De percentage verdelingen, die men zo krijgt zien er voor de totale PO en NO groep als volgt uit:

	PO	NO
Zeër goed	22%	19%
Goed	42%	20%
Slecht	35%	60%
	<hr/>	<hr/>
	100%	100%

Aangezien de getallen, binnen één categorie van zeer goed tot slecht, van PO en NO nu direct vergeleken mogen worden, is nu zonder meer te zien dat de categorie 'zeer goed' binnen PO ongeveer even groot is als binnen NO, doch dat er in PO aanzienlijk meer 'goede' resultaten zijn gevonden terwijl binnen NO veel meer 'slechte' resultaten zijn gescoord.

Nu moet men zich afvragen of de beslissing hiermede is gevallen:

De primaire behandelingsmethode voor enkelfracturen zoals door Weber werd aan-gegeven is duidelijk de voorkeursbehandeling ten opzichte van de oudere meer conservatieve behandelingsmethode.

Indien alle gebruikte parameters een onomstotelijke betekenis hadden, 'zeer goed' altijd onaanvechtbaar beter was dan 'goed' etc., en zeker was dat de duidelijke verdelingsverschillen niet op toeval kunnen berusten dan zou de beslissing inderdaad gerechtvaardigd zijn. Zo eenvoudig is het echter niet.

Een verantwoorde beslissing kan men op twee manieren bereiken, waarbij direct moet worden vermeld dat talloze mengvormen mogelijk en vaak nodig zijn:

1) Men kan de gehele gang van zaken bij het onderzoek nauwkeurig bestuderen, nagaan of men het eens kan zijn met de uitgangs hypothesen, beoordelen of men de diverse redeneringen 'gezond' vindt en tenslotte nagaan of men van mening kan zijn dat het toeval een belangrijke rol kan spelen in de verkregen verdelingsverschillen bij de eindscores. Zo kan men op basis van de verkregen indruk een min of meer gevoelsmatige of intuïtieve beslissing nemen betreffende de betekenis van de onderzoekresultaten. Bij het nemen van een dergelijke beslissing wordt een brede scala van informatie betrokken.

Goed geformuleerde gegevens worden gewikt en gewogen, waarbij min of meer onbewuste statistische processen een rol spelen. Ook nauwelijks geformuleerde gegevens worden in de overwegingen betrokken, terwijl ook onbewust geregistreerde informatie de uiteindelijke beslissing beïnvloedt.

De kwaliteit van een dergelijke gevoelsmatige beslissing kan bijzonder hoog zijn. Het is vrijwel het enige type beslissing dat in de problemen van het dagelijks leven mogelijk is, terwijl alle niet exacte wetenschap tot voor kort vrijwel geheel berustte op de gevoelsmatige beslissing.

De gevoelsmatige beslissing heeft een groot voordeel:

- Ze gebruikt als basis een zeer breed scala van gegevens, dat breder is dan het binnen het onderzoek geformuleerde gegevenspatroon. Daardoor kan de kwaliteit van de beslissing hoger zijn dan strikt genomen, op logische basis uit de gegevens van het onderzoek zelf, mogelijk is.

Ze heeft echter ook een aantal nadelen:

- De beslissingsgronden zijn moeilijk overdraagbaar.
- Daarom is de beslissing altijd aanvechtbaar.

- De beslissing en de beslissingskwaliteit zijn gebonden aan de persoon van degene die beslist.
- De relevantie van de beslissingsgronden is niet te beoordelen en vaak ook aan degene die beslist niet bekend. (De wens is ook de vader van de beslissing!)
- Door de subjectiviteit van de beslissingsgronden zijn talloze beslissingen mogelijk op basis van dezelfde onderzoekgegevens, die echter geen van allen te objectiveren zijn en kan één 'objectief' onderzoek aanleiding geven tot vele tegenstrijdige meningen.
- Kortom de gevoelsmatige beslissing gaat gepaard met een niet geformuleerde onzekerheid over de juistheid ervan.

2) Men kan trachten het onderzoekprobleem onder te brengen in een mathematisch model, dat tot de oplossing van het probleem moet leiden. Zo kan men aansturen op een logische oplossing, die is gebaseerd op een keten van logische beslissingen.

Door de onderlinge ongelijkheid van de proefpersonen, waardoor men slechts zelden in staat is om op grond van metingen verricht bij één proefpersoon generaliseerbare beslissingen te nemen, is men bij mathematische modellen bij geneeskundig onderzoek veelal gebonden aan statistische beslissingen. Immers de genoemde ongelijkheid kan aanleiding geven tot toevallige verschillen tussen de twee te vergelijken groepen gegevens.

Zo worden de parameters van de meeste klinische onderzoeken tot door het toeval beïnvloede (= stochastische) variabelen en het is de mathematische statistiek welke als enige methode in staat stelt, langs reproduceerbare en logisch-mathematische weg stochastische grootheden te beoordelen.

Bovendien stelt de statistische methode in staat de onzekerheid omtrent de juistheid van de beslissing te formuleren. Kansen zijn berekenbaar, waardoor een definieerbaar 'onzekerheidsprincipe' hanteerbaar wordt.

Zo heeft de statistische beslissing een aantal voordelen:

- Zij is reproduceerbaar
- Zij is objectief en slechts op aanwijsbare punten aanvechtbaar. (B.v. de keuze van de significantiegrens).
- De beslissingsgronden zijn bekend.
- De relevantie van de beslissingsgronden is te beoordelen.
- Er is/zijn slechts één of enkele beslissing(en) mogelijk. De gevarieerdheid van meningsverschillen is beperkt.
- De gronden van meningsverschillen zijn formuleerbaar.
- Kortom de statistische beslissing gaat gepaard met een geformuleerde onzekerheid over de juistheid ervan.

De statistische beslissing heeft echter ook nadelen:

- De beslissing kan slechts berusten op gegevens, die werden geformuleerd binnen het onderzoek.
- De kwaliteit van de beslissing hangt af van de kwaliteit van het mathematisch model dat er de grondslag van vormt.
- De beslissing geldt uitsluitend de rol van het toeval in de verdeling van de te beoordelen variabelen. Op welke wijze een toevallige verdeling, of een juist niet toevallig verdelingsverschil, tot stand kwam is een kwestie van interpretatie. De aanvechtbaarheid van deze interpretatie hangt samen met de volledigheid van het model en de hardheid van de meetgegevens.
- Kortom de kwaliteit van de statistische beslissing ten aanzien van een onderzoek-probleem hangt ten nauwste samen met de volledigheid van het gebruikte mathematische model en de kwaliteit van de te beoordelen gegevens, d.w.z. met de kwaliteit van de gehele beoordelingsketen.

Bij het benaderen van de problemen, die binnen een klinisch onderzoek dienen te worden opgelost, beschikt men slechts zelden over een volledig mathematisch model en evenmin over harde en goed te kwantificeren meetgegevens. Men dient zich dan ook te realiseren, dat strikt logische beslissingsketens hier nauwelijks mogelijk zijn. Daartoe is de beschikbare methodologie nog niet toereikend. Op een aantal punten zal men een arbitraire beslissing moeten nemen, die het karakter van een gevoelsmatige beslissing heeft. Dit houdt niet in, dat het opbouwen van een model een vruchteloze zaak is. De gevoelsmatige beslissing wordt localiseerbaar, de statistische eindbeslissing wint aan objectiviteit, de eindinterpretatie verliest aan aanvechtbaarheid en wordt harder. Het statistisch beoordelen van alleen de verdelingsverschillen van een eventuele eindscore, zonder het toetsen van het onderzoekmodel, heeft echter weinig waarde.

De eindconclusie behoudt alle eigenschappen van de gevoelsmatige conclusie en interpretatie en de statistische eindtoetsing verleent er alleen een niet ongevaarlijke schijnbare exactheid aan.

De statistische beslissing is steeds gebaseerd op het beoordelen van incomplete informatie, verkregen uit een aselechte steekproef. De informatie is incompleet ten aanzien van de populatie, waaruit de steekproef werd getrokken en waarvoor de beslissing geldig wordt geacht. Ik ga in hoofdstuk III nader op de principes van de statistische beslissing in. Op deze plaats wil ik volstaan met het invoeren van enkele begrippen, die gehanteerd worden bij het uitvoeren van een statistische toetsing van een verdelingsverschil en die bij de verdere beschrijving van de onderzoekresultaten nog vaak terug zullen komen.

Men denke zich een pot met ondoorzichtige wand, die gevuld is met rode, groene en blauwe knikkers. Men neemt steeds een knikker uit de pot zonder te zoeken en registreert welke kleur men trok. Vervolgens legt men de knikker weer in de pot terug, schudt goed en neemt weer een knikker enz. Men gaat net zolang door tot men b.v. 50 knikkers heeft geregistreerd.

Men heeft nu een *steekproef* met teruglegging van 50 elementen verkregen. De knikkers in de pot vormen de *populatie* en iedere knikker had bij elke trekking steeds weer dezelfde kans om getrokken te worden. Men beoordeelt de fracties rode, groene en blauwe knikkers van de steekproef en bepaalt zo de *verdeling* van de drie kleuren binnen de steekproef.

Men trekt nu een nieuwe steekproef. Nu zal men redelijkerwijze een andere verdeling van de drie kleuren vinden binnen de nieuwe steekproef.

Trekt men op deze wijze een groot aantal steekproeven, dan zal blijken, dat sommige verdelingen van de drie kleuren binnen de steekproef herhaaldelijk optreden, terwijl andere zich slechts zelden voordoen. Toch komen al deze steekproeven uit dezelfde populatie. De kans op het optreden van de verschillende verdelingen is blijkbaar niet gelijk. Op grond van deze proeven kan men een lijst opstellen van de verschillende verdelingen die zijn opgetreden met hun kans om getrokken te worden uit steeds dezelfde populatie.

Dit vormt de basis van een statistische beslissing:

Nu wil men twee verschillende steekproeven vergelijken. Men bepaalt de verdelingen en zoekt in een lijst op hoe groot de kans is op het gevonden resultaat bij trekking uit identieke populaties om er rekening mee te houden dat slechts het toeval het verdelings verschil heeft veroorzaakt. Deze kans wordt de *overschrijdingskans* genoemd. Is die kans groot, dan is weliswaar niet bewezen dat de twee steekproeven uit identieke populaties kwamen, doch de kans is berekend en blijkbaar groot. Is de kans daarentegen erg klein dan heeft men aanleiding om te beslissen, dat de steekproeven uit verschillende populaties kwamen. Hoe klein moet die kans zijn om een dergelijke beslissing te rechtvaardigen?

Hier moet een arbitraire afspraak worden gemaakt. Er dreigen namelijk zowel foute uitspraken bij het aannemen van een te kleine drempel als bij het hanteren van een te grote drempel tot de uitspraak dat niet identieke populaties de oorsprong waren van de vergeleken steekproeven. Indien deze drempel, die *significantiegrens* of *onbetrouwbaarheidsdrempel* wordt genoemd, te groot wordt gekozen dan zal men herhaaldelijk ten onrechte betekenis aan een opgetreden verdelingsverschil hechten, echter bij een zeer kleine betrouwbaarheidsdrempel zal men in een aantal gevallen ten onrechte geen betekenis hechten aan een verdelingsverschil. Deze twee soorten fouten worden *fouten van de eerste- respectievelijk van de tweede soort* genoemd. De overschrijdingskans, die veelal wordt aangeduid met de letter P, wordt bij veel onderzoeken wanneer  $P < 0,05$  als significant klein beschouwd, d.w.z. de onbetrouwbaarheidsdrempel wordt vaak op 0,05 gesteld. Men dient zich hierbij te realiseren, dat bij beslissingen op grond van deze drempel in ongeveer 1 op 20 beslissingen een fout van de eerste soort kan worden gemaakt.

De overschrijdingskans P is meestal niet eenvoudig direct te berekenen op basis van de verdeling binnen een steekproef. In het beschreven voorbeeld bestonden de steekproeven uit de kleuren van de getrokken knikkers. Meestal bestaan steek-

proeven uit variabelen, die min of meer gekwantificeerd zijn geregistreerd. Een groep dergelijke cijfers kan men beoordelen op basis van hun gemiddelde, mediaan, standaardafwijking of andere, de groep beschrijvende, *kengetallen*. Twee steekproeven kan men dan ook vergelijken op grond van verschillen, of overeenkomsten van hun kengetallen. Op basis van deze kengetallen zijn talloze statistische toetsingsgrootheden ontwikkeld. Met behulp van een formule kan men uitgaande van een of meer kengetallen van de te vergelijken steekproeven de waarde van de toetsingsgrootheid berekenen. Door middel van tabellen is uit deze waarde de gezochte overschrijdingskans terug te vinden. In het algemeen geldt dat hoe meer kengetallen in de berekening worden betrokken hoe onderscheidender een toetsingstechniek is, aangezien dan meer karakteristieken van de steekproef bij het berekenen van de toetsingsgrootheid worden betrokken. Een onderscheidende toets is in staat bij kleinere verschillen tussen steekproeven een kleine overschrijdingskans te geven, dan een minder sterk onderscheidende. Het is dus van belang in iedere situatie naar een passende toetsingstechniek te zoeken. Iedere toets heeft een beperkt toepassingsgebied en eigen karakteristieken.

Er wordt steeds gesproken over een toetsing, doch er is nog niet gesproken over hetgene dat eigenlijk wordt getoetst. In feite wordt meestal de hypothese getoetst dat de te vergelijken steekproeven uit identieke populaties werden getrokken, als alternatieve hypothese wordt dan gesteld, dat de steekproeven uit niet identieke populaties afkomstig zijn. Is de berekende overschrijdingskans kleiner of gelijk aan de afgesproken significantiegrens dan verwerpen we de z.g. *nulhypothese* ten gunste van de *alternatieve hypothese*. Deze nulhypothese en eventuele alternatieve hypothesen spelen een centrale rol bij de statistische beslissing.

Ze bepalen de richting van de beslissing, hebben invloed op de overschrijdingskans en het is dan ook zaak om bij iedere statistische beslissing nauwkeurig de nulhypothese en de alternatieve hypothese op te stellen. Ze worden als regel aangegeven met  $H_0$ , resp., met  $H_1$ .

### **C Resultaten van het onderzoek en analyse van de bevindingen**

In tabel 1 zijn de onderzoekresultaten samengevat naar de beoordelingscriteria van Weber. In de bovenste rij zijn de totaalresultaten te vinden, terwijl de drie rijen eronder de resultaten van de A-, B- en C-fracturen uit groep luxatiefracturen weergeven.

In de linker kolom zijn steeds de betrokken cijfers van Weber vermeld, terwijl de twee rechter kolommen de uitkomsten van de PO- en de NO groep weergegeven.

In tabel 2 zijn de uitkomsten weergegeven van statistisch getoetste vergelijkingen. Gebruikt werd de Mann Whitney U toets. De waarde van de normaal verdeelde toetsingsgrootheid  $z$  is in de meest rechter kolom aangegeven. De overschrijdings-



kansen, die kleiner zijn dan de significantiegrens 0,05 zijn met een \* aangegeven. Welke steekproeven met elkaar werden vergeleken, wordt in het linker gedeelte aangegeven.

Uit de tabellen blijkt, dat bij vergelijking van de twee steekproeven PO en NO de  $H_0$  (de steekproeven werden getrokken uit identieke populaties) ten gunste van de alternatieve hypothese moet worden verworpen. Het aantal zeer goede resultaten is bij de geopereerde groep niet zoveel beter, doch de winst zit vooral in een belangrijk groter aantal goede resultaten in de PO groep.

Selecteren we vervolgens uit de PO en NO groepen de gegevens afkomstig van de A-, B- en C-fracturen uit de groep luxatiefracturen, dan blijkt uit de tabellen het volgende:

Vergelijking van de A-fracturen uit de PO groep met die uit de NO groep levert geen aanleiding om  $H_0$  te verwerpen ( $P=0,18$ ).

Vergelijking van de B- en de C-fracturen uit de PO groep met dezelfde subgroepen uit de NO groep levert steeds een significant verschil op ( $P$  resp.  $<0,0007$  en  $<0,004$ ).

Bij beide groepen blijkt de winst wederom voornamelijk in het aantal 'goede' resultaten te zitten.

Het lijkt dus alsof bij de geopereerde patiënten, althans met een B- en C- fractuur, inderdaad een beter resultaat werd gevonden dan bij de niet operatief behandelde patiënten met vergelijkbare fracturen. Doch zoals reeds op pag. 34 werd gesteld heeft toetsing van alleen een eindscore weinig betekenis.

Weber stelde, dat bij niet operatieve behandeling, de A-fracturen een betere prognose hebben, dan de B-fracturen. Als we deze bewering trachten te toetsen en deze subgroepen uit de NO groep selecteren, dan blijken we de nulhypothese, dat de steekproeven uit identieke populaties werden getrokken, te moeten verwerpen:  $P=0,006$ . Bij vergelijkbare groepen uit de PO groep blijkt echter geen verschil meer aantoonbaar te zijn. ( $P=0,21$ ).

Tot zover kunnen we steeds een bevestiging van de stellingen van Weber vinden. Hij stelde echter eveneens, dat bij niet operatieve behandeling de prognose van B-fracturen beter was, dan die van C-fracturen en dat dit prognoseverschil bij operatieve behandeling verdween. Toetsing van geschikte subgroepen om deze bewering te toetsen leidde echter noch bij de gegevens uit de NO groep, noch bij die uit de PO groep tot verwerping van  $H_0$ . M.a.w. verschil in prognose tussen B- en C- fracturen was noch bij de niet operatief behandelde-, noch bij de operatief behandelde groep aantoonbaar. ( $P$  resp. 0,11 en 0,14). Hier kunnen we dus geen ondersteuning van de opvatting van Weber vinden.

Men dient zich nu te gaan afvragen, of de binnen dit onderzoek gevonden resultaten verschil vertonen met de door Weber gepubliceerde gegevens. Hiertoe werden zowel de totale PO groep, als de subgroepen samengesteld uit A-, B- en C-fracturen uit de PO groep met vergelijkbare groepen van Weber vergeleken.

Het blijkt (zie ook tabel 2):

Weber — PO  $P < 0,00003$  (Totale groepen)

Weber — PO  $P = 0,01$  (A-fracturen)

Weber — PO  $P < 0,00003$  (B-fracturen)

Weber — PO  $P < 0,00003$  (C-fracturen)

De resultaten van de operatieve behandeling zoals ze door Weber werden vermeld zijn dus in alle groepen significant beter dan de resultaten, die voor de patiënten uit de PO groep werden geregistreerd. Het blijkt dus dat de PO groep weliswaar een beter behandelingsresultaat vertoonde dan de conservatief behandelde NO groep, doch dat de resultaten van Weber, met in principe dezelfde behandeling als bij de patiënten uit de PO groep weer veel beter uitvielen dan de PO groep. Deze bevinding eist een nader onderzoek: verdere vergelijking tussen PO groep en W-groep brengt aan het licht, dat het tijdsverloop tussen ongeval en naonderzoek van de patiënten uit de PO groep gemiddeld veel langer is dan bij de patiënten uit de W groep.

Bij de PO groep is dat tijdsverloop gemiddeld 180 weken, terwijl het bij het onderzoek van Weber slechts 61 weken bedraagt.

Om te toetsen of dit verschil meer dan toevallige betekenis heeft, zouden meer gegevens nodig zijn dan Weber in zijn boek geeft. Het verschil in gemiddelde naonderzoekperiode is echter zo groot dat het niet al te gewaagd lijkt te zijn, om te veronderstellen dat een eventueel opgestelde nulhypothese verworpen zou moeten worden.

Indien echter het verschil in tijd de verklaring zou moeten geven, voor het feit dat de resultaten van het eigen onderzoek zo sterk verschillen met de door Weber gepubliceerde resultaten, dan moet tegelijk worden geconcludeerd, dat de eigen resultaatcijfers niet vergelijkbaar zijn met die van Weber. Men zou immers kunnen veronderstellen, dat indien de naonderzoekperiodes bij beide steekproeven (W en PO) wel gelijk waren geweest, ook de resultaten identiek waren uitgevallen.

Aangezien we over geen enkel gegeven beschikken, dat ons tot een ontkenning of bevestiging van deze veronderstelling kan leiden, moeten we tot de conclusie komen dat het gevonden resultaatverschil tussen de W en PO steekproeven niet te interpreteren is. De gegevens van de twee steekproeven zijn dus onvergelijkbaar.

Men zou kunnen proberen een correctiefactor voor het verschil in tijdsverloop tussen fractuur en naonderzoek (naonderzoekperiode) te bepalen, waarmee dan wel vergelijkbare eindcijfers zouden moeten ontstaan. Dit geeft aanleiding om de gehele weg waarlangs het (eind)resultaatcijfer werd bepaald, nader onder de loep te nemen.

Het probleem, waar we nu mee te maken hebben gekregen, n.l. het moeilijk vergelijkbaar zijn van de resultaten van twee op hetzelfde doel gerichte onderzoeken komt veel vaker voor en waar het wordt gesignaleerd levert het steeds problemen op voor de interpretatie van een onderzoekresultaat.

Souquet werd reeds op pag. 14 geciteerd, Sullivan, Watkins, Oberfield en Khazel schreven in een artikel over 'arterial infusion cancer chemotherapy' in 1967: 'Moreover, it is virtually impossible to assess and compare the results obtained by these investigators because of the lack of standardization of methods of treatment reported by different investigators in relation to the following factors: surgical techniques of catheter introduction, types of apparatus used, conduct of infusion, drugs and drug doses schedules employed, duration of treatment, selection of patiënts, and statistical evaluation of results.....'

Tallos velen van dit soort citaten zijn in de klinische literatuur te vinden. Meestal is het gebruik van niet identieke klassificaties de meest opvallende oorzaak van de onmogelijkheid tot vergelijking. Sullivan c.s. noemen echter bovendien het gebrek aan standaards, selectie van patiënten en de statistische bewerking als een oorzaak van de moeilijke vergelijkbaarheid van de onderzoekresultaten, hoewel het niet helemaal duidelijk wordt wat ze daar eigenlijk mee bedoelen. Ook Weber gebruikt een geheel eigen klassificatie van de letsels van het B.S.G. Er dient echter te worden vermeld, dat sinds hij de klassificatie introduceerde, deze binnen de A.O. algemeen wordt gebruikt.

Juist met het oog op te bereiken vergelijkbaarheid werden de fracturen uit de groepen PO en NO volgens de Weber klassificatie ingedeeld. Toch blijkt er nu een onvergelijkbaarheid te zijn opgetreden, doch door een geheel andere reden. Dit geeft aanleiding om te veronderstellen dat er nog meerdere oorzaken voor het optreden van onvergelijkbaarheid kunnen zijn. Dit was een reden om tot een zo nauwkeurig mogelijke analyse te komen van alle factoren, die aanleiding kunnen geven tot onduidelijkheden van de eindresultaten en zo de vergelijking met andere onderzoekresultaten kunnen bemoeilijken. Hierbij kwamen een aantal oorzaken naar voren. Deze bleken in hoofdzaak terug te brengen te zijn tot een zestal foutenbronnen. Zonder het elimineren van deze foutenbronnen lijkt de vrees gerechtvaardigd dat de betekenis, de vergelijkbaarheid en de nauwkeurigheid van klinisch naonderzoek steeds onduidelijker zal worden. Het elimineren van deze foutbronnen zal grote problemen opleveren. Het oplossen van deze problemen lijkt echter van zo fundamentele betekenis, dat een bespreking van deze problemen en een analyse van de mogelijkheden tot het vinden van oplossingen niet achterwege mag blijven.

*There is nothing more necessary to the man of science than its history, and the logic of discovery . . . : the way error is detected, the use of hypothesis, of imagination, the mode of testing.*

LORD ACTON

### Hoofdstuk III

## FOUTENBRONNEN BIJ KLINISCH NAONDERZOEK

De te bespreken foutenbronnen zijn frequent in de literatuur te signaleren. Door onvoldoend scherpe probleemformuleringen, het niet ter beschikking zijn van standaards en vaak onbewust gebruik van niet getoetste hypothesen ontstaat een onderzoekmodel dat weinig bescherming biedt tegen het binnensluipen van onnauwkeurigheden, die niet of nauwelijks worden opgemerkt. Te vaak wordt een onderzoek verricht op basis van vage vraagstellingen waardoor de bewerkingsmethodiek pas na het verzamelen van de meetgegevens moet worden geformuleerd. Soms blijkt dan te laat, dat de meettechniek niet optimaal is geweest en zijn de onderzoekresultaten minder hard, dan ze hadden kunnen zijn indien het onderzoekmodel voor de metingen was geformuleerd.

De besproken foutenbronnen worden alle signaleerd binnen de methode, die Weber toepaste bij de bewerking van zijn naonderzoekgegevens. Er zij hier met klem op gewezen, dat dit alleen gebeurt omdat slechts de analyse van zijn methode binnen het bestek van dit onderzoek past. De signalering zou even goed naar aanleiding van andere publicaties kunnen plaats vinden. Het is juist de volledigheid waarmee Weber zijn onderzoek beschrijft, die de gelegenheid opent om de fouten op te sporen.

Globaal gesproken helpen alle te bespreken fouten mee om het onderzoekmodel kwalitatief te verminderen. Hierdoor lijdt de hardheid van de onderzoekconclusie schade, waardoor ze aan bewijskracht inboet. De interpretatie ervan krijgt dan steeds meer een gevoelsmatig karakter waardoor het oorspronkelijke onderzoekprobleem niet wezenlijk is opgelost, doch slechts een nieuwe bron voor meningsverschillen is geschapen.

De zes te bespreken fouten zijn:

- A. Onvergelijkbaarheidsintroductie door gebruik van ongelijke klassificaties.
- B. Verkeerd gebruik van hypothesen.
- C. Onduidelijke inhoud van de beoordelingsnorm.
- D. Fouten bij het kwantificeren van waarnemingen.

E. Verwaarlozing van de factor tijd.

F. Verwaarlozing van de invloed van selectie bij het vergelijken van klinische steekproeven.

#### **A** *Onvergelijkbaarheidsintroductie door gebruik van ongelijke klassificaties.*

Deze eerste foutenbron wordt regelmatig in allerlei klinische literatuur gesignaleerd en reeds op pag. 14 en pag. 39 besproken. Het is merkwaardig, dat dit probleem tegelijk zo bekend is, en tevens zo vaak kan worden signaleerd. Het inzicht, dat een onderzoek alleen maar aan waarde wint door het streven naar een zo groot mogelijke vergelijkbaarheid met relevante andere onderzoeken lijkt nog niet overal aanwezig te zijn. Zelfs de introductie van een 'betere' klassificatie gaat altijd gepaard met onvergelijkbaarheid van het onderzoek, tenzij tevens een meer standaard indeling wordt gevolgd. Het 'betere' is echter altijd maar de vraag en de waardevermindering door de onvergelijkbaarheid een feit!

#### **B** *Verkeerd gebruik van hypothesen*

Onjuist gebruik van hypothesen kan de duidelijkheid en de hardheid van het onderzoek ernstig benadelen. Twee typen fouten moeten hierbij worden onderscheiden:

- Gebruik van een hypothese op een verkeerde plaats in de beoordelingsketen.
- Hanteren van niet-getoetste hypothesen alsof het bewezen feiten zijn.

De eerste fout wordt door Weber gemaakt doordat hij de hypothese waarop zijn behandelingstechniek berust, tevens gebruikt bij het kwantificeren van de naonderzoekresultaten. Zo sluit hij als het ware zijn onderzoek kort en zal altijd bewijzen dat zijn behandelingstechniek goed is!

In hoofdstuk I bleek reeds dat de behandelingsprincipes van Weber berusten op interpretaties van anatomische en functionele bijzonderheden van de structuren rond het B.S.G. Hij zette deze bijzonderheden om in de bewering: De belasting van het laterale complex: malleolus lateralis+syndesmoses+tibia is gedurende het normale gaan reeds groot, dus nauwkeurige (operatieve) verzorging van dit laterale complex en herstel van de stabiliteit in dit gebied is een absolute vereiste bij het behandelen van fracturen van het B.S.G.

Een dergelijke apodictische bewering berust in feite op de hypothese: Niet operatieve behandeling van letsels in het gebied van het laterale complex van het B.S.G. geeft een minder goed resultaat dan operatieve behandeling van deze letsels. Weber doet nu een onderzoek om na te gaan of zijn opvatting juist is. Hij gaat dus deze hypothese toetsen; hierbij zal hij moeten formuleren wat hij een goed dan wel een slecht behandelingsresultaat noemt. Bij deze formulering mag hij zijn te toetsen hypothese niet gebruiken, doch er moet een van de hypothese onafhankelijke

beoordelingsnorm worden gebruikt. Op basis van zijn behandelingsprincipes worden de letsels van het B.S.G. alle nauwkeurig operatief hersteld en de instabiele banden gehecht. Redelijkerwijze mag nu worden verwacht dat röntgenfoto's na enige tijd genomen minder afwijkingen in het laterale gebied vertonen, dan bij niet operatief behandelde patiënten het geval zou zijn. Daarmee is echter de hypothese niet getoetst aangezien het resultaat van een behandeling niet zozeer van een röntgenfoto afhangt als wel van het subjectieve welbevinden van de patiënt. Bij de kwantificering van de röntgenbevindingen bij het naonderzoek gebruikt Weber echter zijn te toetsen hypothese wel en noemt een onregelmatigheid aan laterale kant slecht. Hij kwantificeert dit verschijnsel met een drie, terwijl hij in de eindscore alle patiënten met een drie of meer als slecht klassificeert. Nu heeft hij zijn onderzoek kortgesloten en kan van tevoren weten dat het resultaat van zijn behandelingsprincipes minder vaak 'slecht' zal uitvallen dan van niet operatieve behandelings technieken. Het zal dan ook later blijken, dat de eindscore van Weber op onaanvaardbare wijze met de kwantificering van de röntgenologische resultaten samenhangt (zie hoofdstuk V). Een resultaatbeoordeling, die van de te toetsen hypothese onafhankelijk was had deze fout kunnen voorkomen.

Het is te betreuren, dat Weber deze fout maakt. Hij doet dat waarschijnlijk omdat hij zijn hypothese zo plausibel vindt, dat hij hem onbewust als een vaststaand feit gaat hanteren. Toch vergist hij zich hierin. Het is niet moeilijk op grond van andere bijzonderheden van het B.S.G. precies de tegengestelde hypothese op te stellen; b.v. op grond van de volgende argumenten:

- Bij het ontstaan van posttraumatische arthrose gaat het niet alleen om de totale kraakbeenbelasting, maar waarschijnlijk veeleer om de belasting per oppervlakte eenheid. Aangezien het gewrichtsoppervlak tussen laterale malleolus en talus een groter oppervlak heeft dan dat tussen mediale malleolus en talus is het niet zonder meer evident, dat de grotere totale belasting aan de laterale zijde gepaard gaat met een grotere belasting per oppervlakte eenheid. Daarmee is het ook niet zonder meer duidelijk dat discongruenties aan laterale zijde een grotere kans op posttraumatische arthrose geven, dan dezelfde discongruenties aan mediale zijde.
- Het opvangen van discongruenties van de benige elementen van een gewrichtsoppervlak gebeurt door het bekledend kraakbeen. Het is voldoende bekend dat onregelmatigheden van de botoppervlakken door regenererend kraakbeen kunnen worden opgevuld, waaruit toch een fraai functionerend gewrichtsoppervlak resulteert. Mede door dit mechanisme kan b.v. de conservatieve behandeling van tibiaplateaufracturen tot een goed functioneel herstel leiden. Locaal operatief ingrijpen zou dit compensatoir mechanisme nadelig kunnen beïnvloeden en theoretisch is dit dus slechts gerechtvaardigd indien aangenomen moet worden dat de benige discongruentie zo groot is dat een decompensatie van het natuurlijke herstelmechanisme is te verwachten.

- De capaciteit van het lokale compensatoire herstelmechanisme zal naar alle waarschijnlijkheid samenhangen met de toestand van het voedend vaatbed. Het is opvallend, dat het vaatpatroon van de mediale malleolus een veel schraler indruk maakt, dan dat van de laterale malleolus (Crock & Crock 1967).

Op grond van de boven genoemde punten is het nu niet vergezocht om de volgende hypothese op te stellen:

De (kleinere) belasting aan de mediale zijde van het B.S.G. in combinatie met het kleinere gewrichtsvlak en het matig regenerend kraakbeen door een armer voedend vaatbed schept omstandigheden waardoor kleine discrepanties van de benige elementen van de mediale malleolus eerder aanleiding geven tot het ontstaan van arthrotische veranderingen van het B.S.G. dan dezelfde discrepanties aan de laterale zijde.

Hoewel een hypothese zonder nauwgezette toetsing nooit als juist aanvaard mag worden, hoe plausibel hij ook klinkt, zijn er daarnaast dus nog voldoende gronden om de hypothese van Weber onzeker te noemen. Waarschijnlijk had Weber deze fout nooit gemaakt indien hij zich bij de opzet van zijn onderzoek verplicht had gevoeld de bij het onderzoek te toetsen hypothese duidelijk te formuleren. Dit leidt tot het op pag. 41 genoemde tweede type fouten in verband met gebruik van hypothesen: hanteren van hypothesen alsof het bewezen feiten zijn.

Hoewel een dergelijk gebruik van niet getoetste hypothesen de bewijskracht van een onderzoek ernstig ondergraaft, kan men er bij klinisch wetenschappelijk onderzoek niet geheel omheen. Hypothesen berusten op inzicht. Ze hebben veel van de eigenschappen van de gevoelsmatige beslissing (zie pag. 32), hun juistheid blijft onzeker, deze onzekerheid ongeformuleerd.

Toetsen van klinische hypothesen is niet altijd mogelijk en was gedurende vele eeuwen vrijwel onmogelijk. Ethische motieven spelen (en speelden) hierbij een rol. Terwijl de statistiek, die het langs logische weg beoordelen van stochastische variabelen mogelijk heeft gemaakt, zich pas gedurende de laatste eeuw heeft ontwikkeld, groeide de geneeskunde in de eeuwen daarvoor reeds uit tot een ingewikkelde keten van hypothesen en gevoelsmatige beslissingen. Berustend op die hypothesen ontstond een therapeutisch arsenaal. Hoe langer een therapeutisch dogma heeft bestaan, des te moeilijker is het om de erachter schuil gaande hypothese nog betrouwbaar te toetsen: hoe kan men de patiënten een zo fundamenteel behandelingsprincipe onthouden? Zo leeft de hypothese ongetoetst voort, hopelijk berustend op een juist inzicht.

Kan men om het gebruik van ongetoetste hypothesen in de klinische praktijk van alle dag niet heen, het blijft verstandig om zich hiervan voortdurend bewust te zijn. Het gevaar dreigt anders, dat het verschil tussen hypothese en getoetste kennis vervaagt, een verschijnsel dat Weber tot het maken van een principiële fout heeft gebracht.

Een wetenschappelijk onderzoek heeft echter juist veelal als doel het omzetten van één of meerdere hypothesen in getoetste kennis. Hierbij is het noodzakelijk de gebruikte hypothesen te formuleren en in de redeneringen alleen getoetste beweringen op te nemen. Is het voortredeneren op hypothesen niet geheel te vermijden dan dient dat toch zoveel mogelijk te worden beperkt en de ingebouwde hypothesen dienen duidelijk signaleerbaar te zijn geformuleerd. Berust de rest van de beslissingsketen die tot de onderzoekconclusie leidt op logische stappen, dan is de conclusie alleen aanvechtbaar op basis van twijfel aan de ingebouwde hypothesen.

Onbewust gebruik van hypothesen kan echter de betrouwbaarheid en de richting van de conclusie van een onderzoek sterk en ongemerkt beïnvloeden.

In dit verband lijkt het zinvol om een aantal van de door Weber gebruikte ongetoetste hypothesen op te sommen:

— Hypothesen afkomstig van anatomische en functionele beschouwingen.

- 1) Door de grote belasting gedurende het gaan zal reeds geringe vermindering van de stabiliteit binnen het laterale complex aanleiding kunnen geven tot verschuivingen tussen de gewrichtsvlakken onderling.
- 2) De verschuivingen, die bij verminderde laterale stabiliteit optreden hebben een vergrote kans op het optreden van posttraumatische arthrose tot gevolg.
- 3) Insufficiëntie van de voorste syndesmose geeft voldoende stabiliteitsverlies om de bovengenoemde verschuivingen te doen optreden.
- 4) Door de grote belasting van het laterale complex geven kleine laterale discongruenties, die bij de behandeling van een fractuur van het B.S.G. niet worden gecorrigeerd, een sterke stijging van de kans op het ontstaan van posttraumatische arthrosis.
- 5) Het nadelig effect van kleine discongruenties van de gewrichtsvlakken aan de laterale malleolus is sterker dan dat van dezelfde discongruenties aan de mediale malleolus.
- 6) Verminderde stabiliteit binnen het laterale complex van het B.S.G. geeft vroeger of later aanleiding tot belangrijke klachten van de patiënt.
- 7) Geringe discongruenties aan de gewrichtsvlakken van de laterale zijde van het B.S.G. geven vroeger of later aanleiding tot klachten van de patiënt.
- 8) Dezelfde discongruenties als onder 7) maar aan mediale zijde van het B.S.G. geven minder aanleiding tot klachten bij de patiënt dan die aan laterale zijde.

— Hypothesen, die voor een deel op de eerste groep steunen, doch meer direct zijn verwerkt in de operatieindicaties. De tekst, die Weber gebruikt, is vertaald en tussen aanhalingstekens geplaatst, de hypothesen zijn door genummerd.



'Iedere beschadiging van het B.S.G., die tot één of meerdere bandletsels of fracturen heeft geleid, is voor ons een absolute indicatie tot operatie'.

9) De niet operatieve behandeling geeft bij de hier genoemde beschadigingen steeds een kleinere kans op een goed resultaat, dan de operatieve behandeling. 'Op iedere leeftijd van de patiënt, is het noodzakelijk anatomische reconstructie te verkrijgen'.

Hier worden twee hypothesen door elkaar gebruikt:

10) Iedere niet anatomische stand na een letsel van het B.S.G. geeft een kleinere kans op een goed herstel, dan een operatief verkregen reconstructie...

en

11) Het gestelde onder 10) is onafhankelijk van de leeftijd van de patiënt. 'Gecompliceerde fracturen worden zonder uitzondering onmiddellijk geopereerd'. Ook hier weer twee te onderscheiden hypothesen, waarvan de eerste gelijk is aan hypothese nummer 9:

12) (= 9) (Gecompliceerde) fracturen van het B.S.G. hebben: bij een niet operatieve behandeling een kleinere kans op een goed herstel, dan bij een operatieve behandeling...

en

13) Uitgesteld opereren van een gecompliceerde fractuur van het B.S.G. geeft een belangrijke verlaging van de kans op een goed herstel.

'Nooit is het tijdstip van een operatie gunstiger dan direct na het ongeval'.

14) Ieder ander tijdstip van operatie ter behandeling van een letsel van het B.S.G. dan direct na het ongeval geeft een verlaging van de kans op een goed herstel.

— Hypothesen, die betrekking hebben op de geadviseerde functionele nabehandeling.

'Vroege mobilisering verbetert de circulatie en het kraakbeen heeft zijn normale glijfunctie nodig, zodat het minder gemakkelijk tot regressieve veranderingen komt'.

15) Na een traumatische beschadiging van het B.S.G. treden onder invloed van gipsbehandeling gemakkelijker regressieve veranderingen in het kraakbeen op dan onder invloed van een functionele nabehandeling, die vaak operatieve behandeling vereist.

'De functionele nabehandeling is het beste middel om posttraumatische dystrophie te voorkomen'.

16) Operatieve behandeling van een letsel van het B.S.G. gevolgd door functionele nabehandeling geeft een kleinere kans op het optreden van post-traumatische dystrophie dan iedere andere vorm van behandelen van een dergelijk letsel, die niet gevolgd wordt door functionele nabehandeling.

Het zou niet moeilijk zijn deze lijst nog aanzienlijk uit te breiden. We beschikken nu eenmaal over slechts weinig betrouwbaar getoetste hypothesen in de praktische geneeskunde.

Binnen een onderzoek zal men echter moeten trachten deze hypothesen zoveel mogelijk op hun juistheid te toetsen. Vooruitlopend op de bewerkingen van hoofdstuk VII is het relevant op te merken, dat bij het eigen onderzoek is gevonden:

- De correlatiecoëfficiënt van onregelmatigheden aan de mediale zijde van het B.S.G. met de subjectieve toetsingsnorm (variabele 178) blijkt ongeveer even groot te zijn als de coëfficiënt tussen dezelfde toetsingsnorm en onregelmatigheden aan laterale zijde. De coëfficiënten zijn respectievelijk .141 en .174, ze zijn daarmee nog laag ook! Dezelfde bevinding geldt deze coëfficiënten als ze berekend worden t.o.v. de toetsingsnorm, die de röntgenologisch waarneembare afwijkingen weergeeft. De coëfficiënten zijn dan resp. .448 en .487. Ze zijn dan dus wel hoger, maar evenzeer liggen ze binnen dezelfde grootteorde. Dit betekent, dat er van de grote betekenis, die Weber hecht aan laterale onregelmatigheden ten opzichte van onregelmatigheden aan mediale zijde niet veel overblijft! Bovendien blijken ze beide sterker met de röntgenfoto samen te hangen, dan met het subjectieve welbevinden van de patiënten.
- Een te wijde enkelvork, i.e. instabiliteit aan de laterale zijde bleek niet aantoonbaar samen te hangen met de subjectieve somscore!
- Er bleek geen relatie aantoonbaar tussen het interval in tijd tussen ongeval en operatie enerzijds en enige subjectieve of objectieve toetsingsnorm anderzijds.

Hoewel moet worden opgemerkt, dat het aantal bij het naonderzoek gevonden afwijkingen in het algemeen relatief laag was, waardoor het onderscheidend vermogen van de gebruikte statistische toetsingen matig is, zijn deze bevindingen voldoende om aan te geven dat de hypothesen van Weber, hoe plausibel ze soms ook klinken, nog bepaald niet bewezen zijn. Het is dan ook van belang dat ze als hypothese gebruikt worden en niet door het veelvuldig hanteren ongemerkt overgaan in beweringen, ervaren als vaststaande feiten.

Deze gang van zaken is nogal eens te signaleren. Apodictisch gestelde hypothesen gaan zo over in klinische regels waardoor het verschil tussen hypothese en feit op de achtergrond raakt.

Deze gang van zaken heeft verscheidene gevolgen, zowel voor de praktische geneeskunde als voor het klinisch onderzoek. Over hetzelfde probleem kunnen meerdere plausibele hypothesen worden opgesteld. Dikwijls zal uit dergelijke hypothesen steeds een ander behandelingsbeleid voortvloeien elk met een zekere mate van succes. De onderlinge resultaatverschillen zullen lang niet altijd zonder nauwkeurige toetsing kunnen worden opgemerkt. Als men een onderzoek naar de betekenis van een dergelijk behandelingsbeleid doet en men hanteert de inmiddels als waar ervaren, maar eigenlijk te toetsen, oorspronkelijke hypothese als basis voor het opstellen van

een kwantificering van het resultaat, dan kan ieder behandelingsbeleid zijn eigen kwaliteit 'bewijzen'. Hiermede kan nooit een uitweg gevonden worden uit de talloze tegenstrijdige meningen, die binnen de klinische literatuur betreffende dezelfde onderwerpen gevonden kunnen worden terwijl het ontstaan van nieuwe tegenstrijdigheden voor de hand ligt. Afgezien van grote problemen betreffende kwantificeringsmethoden en de formulering van toetsingskriteria ligt de fout voor een groot deel aan het gebruiken van een ondeugdelijk onderzoekmodel. Ook de Jonge wijst daarop; in zijn boek over medische statistiek komt hij tot de uitspraak: met het voortbestaan van conclusies uit dit soort onderzoek: '...krijgt de medische literatuur wat ik (de Jonge) noem de kookboekvorm en in stede van de neerslag te vormen van betrouwbare gegevens wordt de therapeutische literatuur tot een soort medische folklore'.

Een goed onderzoekmodel beschermt de onderzoeker tegen het verkeerd hanteren van hypothesen, aangezien binnen het raam van het model niet alleen de probleemstelling van het onderzoek, maar ook de beoordelingsketen goed geformuleerd moet vastliggen en op zijn deugdelijkheid moet worden nagegaan. Hierbij behoren foutief gehanteerde hypothesen tijdigesignaleerd te worden.

Door het gebruik van een ondeugdelijk onderzoekmodel lijdt de hardheid van de eindconclusie van het onderzoek ernstig schade, statistisch toetsen van een dergelijke eindconclusie verandert daar niets aan (zie ook pag. 34).

De geneeskunde ontwikkelt zich snel. Van allerlei kanten bereiken een toenemend aantal nieuwe opvattingen de klinische geneeskunde. Hun waarde ten aanzien van de patiëntenbehandeling moeten ze nog bewijzen. Dit bewijs kan alleen geleverd worden door een klinisch onderzoek. Is hierbij het onderzoekmodel niet goed dan wordt slechts een schijnbewijs geleverd en is weer een bron van tegenstrijdige opvattingen ontstaan, die een even hypothetisch karakter heeft als de oorspronkelijk te toetsen hypothese. Hoewel men bij klinisch onderzoek het gebruik van niet getoetste hypothesen niet geheel kan vermijden, lijkt het aangewezen om een aantal duidelijk geformuleerde eisen aan zo'n onderzoek te stellen:

- Bij klinisch onderzoek dient het gebruik van niet getoetste hypothesen tot het minimum te worden beperkt.
- Waar het gebruik van ongetoetste hypothesen niet te vermijden is, dient dit duidelijk te worden aangegeven.
- Het formuleren van een onderzoekmodel dient aan het verzamelen van de voor het onderzoek benodigde gegevens vooraf te gaan. Hierbij dienen zowel het onderzoekprobleem als de beoordelingsketen duidelijk te worden vastgelegd..

Deze eisen hebben de volgende consequenties voor het bewerken van de gegevens die in het eigen onderzoek betreffende de letsels van het B.S.G. werden verzameld:

- Het onderzoekmodel van Weber, dat als voorbeeld diende voor het onderzoek, blijkt niet te voldoen aan de bovengestelde eisen.

- Met name kan de eindscore van Weber niet worden gezien als een acceptabele weergave van het behandelingsresultaat.
- Naast de door Weber gebruikte gegevens, werden bij het naonderzoek nog een groot aantal andere subjectieve en objectieve gegevens verzameld (zie Hoofdstuk IV).
- Uit deze gegevens dient een beoordelingsscore te worden samengesteld, die onafhankelijk is van de te toetsen hypothesen.
- Indien niet getoetste hypothesen in de beoordelingsketen worden gebruikt dient dat te worden aangegeven.

### **C** *Onduidelijke inhoud van de beoordelingsnorm*

Zoals reeds uit de voorgaande bladzijden bleek, voldoet de eindscore van Weber als beoordelingsnorm van het behandelingsresultaat van enkelfracturen niet. Dit alleen al door het feit dat de binnen het onderzoek te toetsen hypothese er op onaanvaardbare wijze in is verwerkt. Ook de inhoudelijke betekenis van de score is onduidelijk. Hoewel hij pretendeert het behandelingsresultaat weer te geven zodanig dat een onderscheid kan worden gemaakt in zeer goede, goede en slechte resultaten blijkt het slechts om een min of meer willekeurige afknopping te gaan van de somscore die Weber uit zijn zes beoordelingscriteria samenstelt (zie pag. 27 e.v.). Indien bij een naonderzoek b.v. gevonden wordt dat een kleine onregelmatigheid aan de laterale zijde van het B.S.G. op de röntgenfoto te zien is, maar verder geen afwijkingen te vinden zijn scoort de patiënt een drie voor de gesignaleerde afwijking, en zijn somscore wordt daarmee ook drie. Vindt men bij een andere patiënt enorme arthrotische afwijkingen, kan de patiënt niet lopen van voortdurende pijn en zitten zijn spronggewrichten onwrikbaar vast, dan krijgt deze patiënt een hoge somscore. Beide patiënten komen echter gezamenlijk in de categorie 'slecht' van de eindscore van Weber terecht. Het is duidelijk dat deze zelfde eindscore in geen verhouding staat tot het sterke verschil in resultaat tussen deze twee patiënten. Van een reële kwantificering van het eindresultaat is dus geen sprake. Patiënten, die sterk verschillen zowel wat subjectieve klachten, als wat objectief waarneembare afwijkingen betreft kunnen blijkbaar allen dezelfde samenvattende eindscore krijgen. Hiermede wordt de inhoudelijke betekenis van deze score onduidelijk.

Men zou zich kunnen (en moeten!) afvragen, wat de inhoudelijke betekenis van de beoordelingseindscore zou moeten zijn. Dit is afhankelijk van de betekenis, die men aan de onderzoekconclusie wenst te hechten. Indien men het behandelingsresultaat van enkelfracturen wil afmeten aan het meetbaar functieherstel van het B.S.G. dan staat het iedere onderzoeker vrij dat te doen, doch indien deze onderzoeker op basis van de uitkomsten van zijn onderzoek, de conclusie trekt dat een bepaald type behandeling van enkelfracturen blijkbaar de voorkeur verdient dan maakt hij een fout tenzij hij zich bewust is van het feit dat hij bij het behandelen van enkel-

fracturen alleen uit is op het bereiken van een zo goed mogelijk functieherstel van het B.S.G. en niet op een zo goed mogelijk herstelde patiënt. Immers een beter meetbaar functieherstel van het B.S.G. zou wel eens niet altijd samen kunnen gaan met een beter subjectief herstel van de patiënt, hoewel er zonder enige twijfel samenhang bestaat tussen het functioneel herstel van het B.S.G. en het functioneel herstel van de gehele patiënt. Als het mogelijk zou zijn een coëfficiënt op te stellen, die deze samenhang precies weergeeft, zou die zonder twijfel niet de waarde één hebben, die bij een volledige samenhang behoort. Het kleiner zijn dan één betekent dat het hanteren van het functioneel herstel van het B.S.G. als beoordelingsnorm voor het herstel van een patiënt, die werd behandeld voor een letsel van het B.S.G. incomplete informatie geeft:

Indien de onderzoeker nu toch zijn behandelingsbeleid bestuurt met de uitkomsten van zijn onderzoek, dan hanteert hij de niet getoetste hypothese dat de samenhang tussen het functioneel herstel van het B.S.G. en het herstel van de patiënt groot genoeg is om een dergelijke besturing verantwoord te doen zijn. Het is niet onmogelijk dat hij daarin gelijk heeft, het is echter aanvechtbaar en wederom is de onderzoekconclusie hypothetisch. Deze keer is het hypothetisch karakter van de onderzoekuitkomst echter ontstaan door de niet op de toepassing van de onderzoekuitkomst afgestemde, inhoudelijke betekenis van de beoordelingsnorm.

Men kan deze fout vaak signaleren. Zo werd jarenlang de behandelingskwaliteit van maligne afwijkingen afgemeten naar het vijfjaar overlevings percentage. Op zich zelf zijn daar niet zoveel bezwaren tegen. Geleidelijk ging men het echter hanteren alsof het de genezingspercentages betrof. Met andere woorden: sluipenderwijs werd de hypothese gebruikt, dat de kans op vijfjaar overleven zonder tekenen van recidief of metastase identiek was met de genezingskans. Deze hypothese is inmiddels achterhaald, de vijfjaar grens verschoven en op een nieuwe drempel wordt een nieuwe hypothese gebaseerd, zij blijft eveneens ongetoetst en zal zeker weer veranderen.

Ook in dit voorbeeld werd de samenhang tussen een enkel naonderzoekgegeven (vijf jaar overleven) en het genezen van de patiënt overschat. Juist in dit voorbeeld is deze gang van zaken opvallend. Immers maligne afwijkingen geven onbehandeld vrijwel altijd een infauste prognose. De tijd van overleven zelf is daarom de sleutel parameter ten opzichte van de genezingskans en niet een ad hoc opgestelde 5 of 10 jaars periode. Het voordeel hiervan is dat de belangrijkste beoordelingsparameter relatief eenvoudig meetbaar en kwantificeerbaar is. Indien men duidelijk formuleert wat onder genezing (of weer gezond zijn) wordt verstaan, dan is op basis van deze formulering en de overlevingsperiode op niet al te ingewikkelde wijze een algoritme op te stellen waarmee het genezingspunt, althans wat het overleven betreft, voor iedere groep maligniteiten gemakkelijk zou zijn uit te rekenen met behulp van follow-up gegevens.

Een dergelijk genezingspunt heeft een duidelijke inhoud en een bekende relatie met

het genezen van de patiënt, hetgeen van de gehanteerde vijfjaar of tienjaar drempels niet kan worden gezegd.

Indien men met de resultaten van een klinisch onderzoek het behandelingsbeleid wil besturen, dan dient de inhoudelijke betekenis van de beoordelingsnorm, die bij het onderzoek wordt gebruikt, het herstel van de patiënt goed weer te geven.

Iedere andere inhoud geeft het sturende effect van de onderzoekresultaten een hypothetisch karakter.

Voordat men tot een goede formulering van de beoordelingsnorm kan komen zal men een formulering moeten vinden van het herstel van de patiënt. Hierbij mist men node het bestaan van internationaal aanvaarde formuleringen. Indien men echter aanvaardt dat hersteld zijn, genezen en gezond zijn identieke begrippen zijn, dan kan uitgegaan worden van de door de W.H.O. aanvaarde definitie van gezondheid, als uitgangspunt voor de formulering van de genezing. In de W.H.O. definitie worden 'physical, social and mental wellbeing' als voorwaarden voor gezondheid gesteld. Hieruit blijkt onmiddellijk, dat bij vrijwel geen enkel klinisch onderzoek de gezondheid (=genezen zijn) van de patiënten geheel wordt nagegaan. Immers van het sociale en mentale aspect van de gezondheid wordt in verreweg de meeste gevallen hoogstens een vluchtige indruk opgedaan. Men kan dat triviaal vinden in het kader van een naonderzoek van destijds organisch zieken, doch dan hanteert men onmiddellijk (en onbewust) de hypothese dat de afwijking waarvoor destijds werd behandeld, of de situatie die tengevolge van die afwijking en de daarop gevolgde behandeling ontstond, geen storingen in deze facetten van de gezondheid hebben veroorzaakt, die niet tevens in de fysieke gezondheidstoestand tot uitdrukking komen. Dit lijkt een hypothese die onhoudbaar is.

Het bijsturen van het behandelingsbeleid op basis van onderzoeken die beoordelingsnormen hanteren welke uitsluitend het 'physical well-being' als inhoud hebben is dan ook niet ongevaarlijk. Theoretisch kunnen behandelingen ontstaan, die organisch een uitstekend resultaat geven doch sociaal en mentaal grote invaliditeit achterlaten. Gelukkig zal 'inzicht' en 'ervaring' in de vorm van gevoelsmatige beslissingen hier wel grenzen aan stellen. Beter is echter om in het logische onderzoekmodel een beoordelingsnorm te formuleren, die het sociale en mentale aspect van de gezondheid zoveel mogelijk tracht te bevatten.

Een tweede aspect van de formulering van de W.H.O. dat aandacht verdient is het centraal stellen van het subjectieve welbevinden.\*

Op het eerste gezicht is dit voor de hand liggend, het gaat er immers om of men zich gezond voelt. In dat verband is het opvallend dat bij het beoordelen van een naonderzoekresultaat nogal eens een sterk accent ligt op objectief waarneembare afwijkingen. Men gaat er dan blijkbaar van uit dat deze afwijkingen hoog correleren

\* In navolging van Rümke vertaal ik 'well-being' door 'welbevinden' zonder op de omvangrijke discussies, die over de W.H.O. formulering zijn gevoerd, nader in te gaan.

met het subjectieve welbevinden. Dit is echter een hypothese, die meestal niet is getoetst en de werkelijke samenhang is meestal onbekend. In ander verband wees ik hier reeds op (pag. 49). Ook Weber maakt deze fout (pag. 42).

Een derde facet van de W.H.O. formulering, dat bespreking behoeft is het ontbreken van het uitspreken van een verwachting. Hier schiet de formulering naar mijn inzichten duidelijk te kort, als basis voor een beoordelingsnorm. Er wordt gesproken van gezondheid als een 'state of well-being'. Theoretisch zou dat inhouden dat b.v. een maagcarcinoom mits subjectief nog ongemerkt aanwezig en objectief niet aantoonbaar, de gezondheidsnorm niet aantast. Immers de grenzen van het vaststellen van een 'state of physical well-being' zijn identiek aan de grenzen van het onderscheidend vermogen van het klinisch diagnostisch onderzoek. Indien men zich realiseert, dat geen enkel klinisch onderzoek, hoe uitgebreid ook, absolute zekerheid kan geven ten aanzien van het ontbreken van iedere vorm van fysieke afwijking wordt het duidelijk dat de formulering van de W.H.O. niet aansluit op het algemene begrip gezondheid, tenzij men de 'state of physical well-being' interpreteert als een zuiver theoretische, onmeetbare, absolute fysieke gezondheid. Deze laatste interpretatie is echter onpraktisch en maakt de W.H.O. formulering onbruikbaar als basis voor de inhoud van een beoordelingsnorm. Het is dan ook te betreuren dat de W.H.O. formulering onduidelijk blijft over de betekenis van de 'state of well-being'. Een beter hanteerbare formule ontstaat door het meer uit elkaar halen van de subjectieve, en de objectief waarneembare gezondheidstoestand. Hierbij dient de subjectieve gezondheidstoestand in het centrum te staan. Op basis van de objectief waarneembare gezondheidstoestand kan een verwachting worden uitgesproken over de subjectieve toestand in de toekomst. Een dergelijke formulering bestaat dus uit twee aparte onderdelen en zou b.v. kunnen zijn:

Gezondheid is

- een toestand van fysiek, sociaal en mentaal subjectief welbevinden, samen gaand met
- een voor betrokken individu normaal te noemen kans dat de toestand van subjectief welbevinden gedurende een bepaalde tijd zal blijven bestaan.

Het eerste onderdeel van deze formulering is gedurende een onderzoek te bepalen, de kans, die in het tweede onderdeel wordt genoemd zou uit de uitkomsten van het lichamelijk onderzoek kunnen worden berekend. De 'bepaalde periode' zou een standaard periode kunnen zijn, die afhankelijk van leeftijd, geslacht en milieu-omstandigheden in principe uit bevolkingsgegevens te berekenen is, terwijl ook de normen voor de genoemde kans binnen deze gegevens bepaald kunnen worden. Nog vele problemen blijven onopgelost, zoals de keuze van de bevolkingsgegevens waaruit de standaard periode en de normale kans zouden moeten worden berekend, problemen, die ik reken tot het gebied van de klinische methodologie. Zeker is, dat het beste mathematische onderzoekmodel een onderzoekconclusie waarmee een

behandelingsbeleid wordt beïnvloed niet tegen een hypothetisch karakter kan beschermen zolang de inhoudelijke betekenis van de, bij het onderzoek gehanteerde, beoordelingsnorm het herstel, c.q. het genezen zijn, c.q. de gezondheid van de onderzochte patiënten nog niet nauwkeurig weergeeft.

Ten aanzien van het eigen onderzoek naar de resultaten van de behandeling van letsels van het B.S.G. zullen deze beschouwingen de volgende consequenties hebben:

- De nieuw te formuleren beoordelingsnorm zal een subjectief onderdeel hebben. Het betrekken van het sociale en mentale aspect in de beoordeling van het herstel van de patiënten is beperkt tot de gegevens die het naonderzoek opleverde. Er zal moeten worden aangenomen, dat deze facetten tot uiting komen binnen de samen te stellen somscore van parameters die tijdens het naonderzoek werden bepaald. De somscore zal worden geacht het subjectieve welbevinden van de patiënten weer te geven.
- Aangezien de relatie tussen de objectief waarneembare afwijkingen en de prognose omtrent het subjectief welbevinden niet bekend is, zal een somscore worden samengesteld betreffende alle gegevens, die betrekking hebben op objectief waarneembare afwijkingen.

Deze twee somscores zullen beiden als beoordelingsnorm worden gehanteerd, ze worden in hoofdstuk V nog nader besproken.

#### **D *Fouten bij het kwantificeren van waarnemingen***

Bij een onderzoek worden als regel waarnemingen geregistreerd. Hierdoor ontstaan gegevens, die via bewerking tot de onderzoekconclusie moeten leiden. Aan ieder gegeven zijn twee facetten te onderscheiden:

- 1) De inhoudelijke betekenis, of de kwaliteit.
- 2) De sterkte, of de kwantiteit.

De inhoudelijke betekenis van de geregistreeerde waarneming dient duidelijk te zijn, zoals reeds in het vorig onderdeel werd besproken ten aanzien van de beoordelingsnorm.

Ook de sterkte van de geregistreeerde waarneming is voor het bereiken van een betrouwbare onderzoekconclusie belangrijk. Immers het is b.v. een groot verschil of een patiënt meedeelt soms nog wel eens een pijnlijk gevoel in de enkel te hebben, of dat hij vertelt geen nacht te kunnen slapen van de pijn die zelfs in rust optreedt. De pijn (kwaliteit) is in beide gevallen aanwezig, doch de sterkte van de pijn (kwantiteit) is verschillend.

Dit verschil dient in de registratie tot uitdrukking te komen (kwantificeren), zodat het bij de bewerking van de gegevens kan worden betrokken, anders lijdt de nauwkeurigheid van de onderzoekconclusie ernstig schade.



Het kwantificeren van waargenomen eigenschappen kan men op meerdere manieren realiseren:

1) Indien men voor de eigenschap een eenheid kan definiëren en men kan meten hoeveel eenheden er worden waargenomen, dan verkrijgt men een intervalschaal. Deze methode verdient altijd de voorkeur aangezien de waargenomen kwantiteit per definitie en reproduceerbaar vaststaat.

2) Indien een eigenschap niet direct meetbaar is, doch men meent wel te kunnen waarnemen of hij zich in mindere of meerdere mate voordoet, dan kan men een volgorde opstellen. Een lager getal geeft dan een lagere sterkte van de waargenomen eigenschap aan dan een hoger getal (rangnummer). Een verschil in rangnummer geeft nu nog slechts een verschil in sterkte aan, de grootte van het verschil is niet gedefinieerd en niet reproduceerbaar; men heeft de waarneming ondergebracht in een geordend categorisch systeem.

3) Indien de eigenschap zich noch direct laat meten noch waarnemen, dan kan men soms objectieve waarnemingen verrichten bij eigenschappen waarvan men weet dat ze hoog correleren met de te kwantificeren eigenschap.

Op grond van waarnemingen bij de verwante eigenschap kan men een schatting verrichten ten aanzien van de te kwantificeren eigenschap en zo weer tot een volgorde komen.

4) De laatste manier, waarop men nogal eens eigenschappen in getalwaarden ziet uitgedrukt is in feite een foutieve toepassing van de methode van punt 3). Ik zou deze kwantificeringen dan ook 'pseudokwantificeringen' willen noemen.

De te kwantificeren eigenschap laat zich noch direct meten noch waarnemen. De relatie met de 'verwante' parameter, die men voor het opstellen van een volgorde heeft gebruikt is echter zo hypothetisch en aanvechtbaar dat sterke twijfel gerechtvaardigd is of de opgestelde volgorde ook maar enige relatie heeft met de sterkte van de te kwantificeren eigenschap.

Tenslotte moet worden opgemerkt, dat men in een aantal gevallen, waar het kwantificeren grote moeilijkheden geeft, kan kiezen voor het laten vervallen van het aangeven van de sterkte. Men geeft dan slechts aan, dat de betrokken eigenschap al of niet waarneembaar was.

Als we nu terugkeren tot de kwantificeringen, die Weber heeft aangegeven, dan blijkt:

- De mate van pijn wordt uitgedrukt in volgnummers. De verwante eigenschap, die wordt gebruikt, is de subjectieve pijnbeleving van de patiënt bij toenemende belasting van het B.S.G.
- Het lopen wordt in volgnummers gekwantificeerd. De diverse loopkwaliteiten zijn direct waar te nemen.
- De activiteit wordt in volgnummers gekwantificeerd. Bij het opstellen worden echter twee verwante parameters door elkaar gebruikt: enerzijds de

beroepsactiviteit van de patiënt en anderzijds zijn activiteit buiten het beroep. Daardoor wordt de opgestelde volgorde aanvechtbaar. Immers, de onderlinge verhouding tussen belasting van het B.S.G. binnen en buiten het beroep is bij iedere patiënt anders en daarom strikt individueel gebonden. Hiermee wordt het onzeker of een verschil in rangnummer ook een verschil in activiteit vertegenwoordigt en is een pseudokwantificering ontstaan.

- De op de röntgenfoto waargenomen afwijkingen worden gekwantificeerd. Aangezien het hier om een objectief waarneembare afwijking gaat zal het cijfer de basis moeten zijn voor het bepalen van de prognose omtrent het subjectief welbevinden.

In de kwantificering van de röntgenwaarnemingen wordt dus eigenlijk de prognostische betekenis van deze afwijkingen gekwantificeerd. De verwante parameter, die wordt gehanteerd is de röntgenologisch waarneembare arthrose. Deze wordt op zijn beurt afgeleid van een aantal röntgenologische kenmerken.

De mate van aanwezigheid hiervan speelt echter geen rol. Zowel door dit laatste als door de opgestelde volgorde zelf (Zie pag. 28). ontstaat zo'n aanvechtbare relatie met de te kwantificeren parameter dat ook hier sprake is van een pseudo-kwantificering.

- De functie van het B.S.G. wordt gekwantificeerd met behulp van de in hoekgraden meetbare bewegingsuitslag. Weber verwerkt in de kwantificering echter tegelijkertijd een vergelijking met de bewegingsmogelijkheden van het gezonde B.S.G. De onderlinge ongelijkheid verdeelt hij in ongelijke klassen, zodat tenslotte een volgorde ontstaat.
- De bewegingsuitslag van het onderste spronggewricht (O.S.G.) is moeilijk in hoekgraden meetbaar. Na vergelijking met de gezonde zijde stelt Weber een volgorde op.

Het opstellen van pseudokwantificeringen is zonder meer fout. De fout ontstaat echter pas in de bewerking. Een volgorde zoals Weber samenstelt voor de registratie van de op de röntgenfoto waargenomen afwijkingen kan men ook beschouwen als een codering van afwijkingen die werden waargenomen. Men heeft deze waarnemingen dan in een nominaal categorisch systeem ondergebracht. Ze kunnen punt voor punt in hun relaties met andere onderzoeksgegevens worden bekeken waarin ze dan worden beschouwd als eigenschappen die zonder kwantificering werden geregistreerd bij aanwezigheid.

Hanteert men hen in de bewerking echter alsof ze een kwantificering van éénzelfde eigenschap vertegenwoordigen, dan maakt men een fout en de conclusie van het onderzoek wordt onbetrouwbaar. Het is dus meer de vorm van hanteren dan het opstellen zelf dat tot fouten leidt. Zo kunnen eveneens fouten ontstaan indien men bij de gegevensbewerking, in rangorde gekwantificeerde gegevens in algoritmen betreft, die ontwikkeld zijn voor gegevens die met een intervalschaal zijn gekwan-

tificeerd. Hoewel de fouten die hiermede ontstaan meestal kleiner zijn dan in het eerste geval kunnen toch conclusies ontstaan die onvoldoende grond in het onderzoek hebben.

Voor het in dit boek beschreven onderzoek hebben deze beschouwingen de volgende consequenties:

- Zoveel mogelijk is getracht de waarnemingen met een interval schaal te kwantificeren.
- Indien volgordes worden gebruikt, zal steeds de eventueel te gebruiken verwante eigenschap een duidelijke relatie moeten hebben met de te kwantificeren eigenschap.
- Indien geen enkele betrouwbare volgorde kan worden opgesteld, wordt van kwantificering afgezien en alleen het al of niet aanwezig zijn van de eigenschap vastgelegd.

#### **E** *Verwaarlozing van de factor tijd.*

Een klinisch naonderzoek is een speurtocht naar veranderingen in de subjectieve en objectieve gezondheidstoestand van patiënten, sinds het laatste contact dat in de patiëntenstatus is vastgelegd.

Een dergelijk onderzoek wordt dus gedaan in het bewustzijn, dat na verloop van tijd veranderingen kunnen zijn ontstaan.

Van belang zijn alleen veranderingen, die in verband staan met de oorspronkelijke afwijking van de naonderzochte personen.

Na hoeveel tijd treden deze veranderingen niet meer op?

Dat wil zeggen, na hoeveel tijd mag men van een eindtoestand spreken? Hierover is weinig met zekerheid bekend. Op pag. 49 werd de hypothetische genezingsperiode van vijf jaar, die een tijdlang ten aanzien van maligne afwijkingen werd gehanteerd, reeds beproven. Deze periode bleek onhoudbaar.

Posttraumatische arthrose zou in de eerste twee jaar na een letsel ontstaan. Was na twee jaar geen arthrose waarneembaar, dan zou het ook niet meer optreden. Deze opvatting, die nogal eens is te signaleren, is gebaseerd op de ervaring, dat het merendeel der patiënten dat na een letsel klachten houdt of krijgt, binnen twee jaar terug komt ter behandeling. Dit is echter onvoldoende basis voor de uitspraak dat men na een periode van twee jaar na een letsel, van een eindtoestand mag spreken. Bij een letsel van het B.S.G. wordt een aantal verschillende weefsels beschadigd, zoals gewrichtskapsel, ligamenten, bot, kraakbeen, bloedvaten, zenuwvezels. Het lijkt plausibel, dat al deze weefsels zich in meerdere of mindere mate incompleet kunnen herstellen.

Een groot aantal verschillende combinaties van onvolledig herstel is dan ook denkbaar. Naast arthrose lijkt het optreden van een grote variatie van verklevingen en

locale fibrosis dan ook te verwachten. Een aantal van deze variaties kan vroeger of later tot subjectieve klachten aanleiding geven, waarbij het aannemelijk lijkt dat de ene combinatie daar eerder aanleiding toe geeft dan de andere. Op dezelfde gronden zou men kunnen verwachten, dat de ene combinatie sneller tot waarneembare afwijkingen aanleiding geeft dan de andere. Zonder een nauwkeurig onderzoek naar de ontwikkelingssnelheid van deze late complicaties lijkt het dan ook nogal gewaagd om een tijdsdrempel te hanteren, waarbinnen alle complicaties zich tot een eindstadium zouden hebben ontwikkeld. Voorlopig lijkt het noodzakelijk om te veronderstellen dat een patiënt, die na twee jaar wordt onderzocht en in de eindbeoordeling 'goed' scoort, b.v. na zes jaar onderzocht in de 'slechte' categorie zou kunnen vallen.

Dit houdt echter in, dat ten aanzien van de beoordeling van een behandelingsresultaat, een patiënt die twee jaar tevoren werd behandeld, niet vergelijkbaar is met een patiënt, waarbij de behandeling zes jaar tevoren plaats vond.

Met andere woorden: bij een naonderzoek, dat ten doel heeft het effect van een behandelingstechniek te beoordelen, dienen de patiënten steeds te worden onderzocht na een zelfde tijdsverloop tussen behandeling en onderzoek. Dit tijdsinterval, dat verder naonderzoekperiode genoemd zal worden, heeft zeer waarschijnlijk invloed op het bij het naonderzoek gevonden behandelingsresultaat en een onderzoekconclusie, die gebaseerd is op onderzoek bij een groep patiënten, waarvan de naonderzoekperiode onderling verschilt, heeft een onduidelijke betekenis aangezien men niet weet voor welk tijdsinterval tussen behandeling en onderzoek het gevonden resultaat relevant is.

Teneinde patiënten die na een verschillende naonderzoekperiode werden onderzocht toch te kunnen vergelijken, zou men kunnen trachten door middel van correctiefactoren de onderzoekresultaten bij deze patiënten naar éénzelfde naonderzoekperiode terug te rekenen. Dit zou theoretisch wel mogelijk zijn, doch alleen indien de invloed van een veranderende naonderzoekperiode op de verschillende naonderzoekbevindingen nauwkeurig bekend is en in een algoritme kan worden uitgedrukt. Dit is voorlopig niet het geval. Om de invloed van de tijd op de naonderzoekbevindingen te leren kennen zal het nodig zijn om onderzoek te verrichten waarbij de patiënten een aantal malen, steeds na een vast tijdsinterval, worden onderzocht. Bijvoorbeeld na twee, vier en zes jaar. Pas nadat de uitslagen van dergelijk onderzoek bekend zijn, is het mogelijk correctiefactoren, zoals boven bedoeld, te berekenen.

Zoals reeds werd opgemerkt, is de relevantie van een onderzoekconclusie die berust op bevindingen bij patiëntenonderzoek, dat met steeds verschillende naonderzoekperiode werd uitgevoerd, onduidelijk. Het vergelijken van twee dergelijke onderzoekconclusies is in principe niet verantwoord, tenzij de invloed van de naonderzoekperiodes gelijk mag worden gesteld. Voor het betrouwbaar nagaan hiervan is het beschikken over de gemiddelde naonderzoekperiode per

groep onderzochte patiënten onvoldoende. Het opgeven hiervan zonder vermelding van andere kengetallen (zie pag. 36.) b.v. de standaard afwijking heeft dan ook niet zoveel waarde. Soms is het mogelijk, bij voldoende spreiding in de naonderzoekperiode in iedere groep, na te gaan of de invloed van de naonderzoekperiode op de verschillende parameters belangrijk is dan wel verwaarloosd mag worden. Zo kan dan een vergelijking, met in acht nemen van bepaalde interpretatierestricties, plaats vinden, ook indien een groot verschil bestaat tussen de gemiddelde naonderzoekperiodes van twee te vergelijken groepen (Zie hoofdstuk VI). Dit blijft echter een hulpmiddel, vermindert het onderscheidend vermogen van het onderzoek en kan beter worden voorkomen door voor goede vergelijkbaarheid zorg te dragen met het constant houden van de naonderzoekperiode van alle onderzochte patiënten.

Binnen het eigen onderzoek waren drie steekproeven te vergelijken.

- de door Weber beschreven groep (W), met een gemiddelde naonderzoekperiode van 61 weken, standaard afwijking niet vermeld, totaal 150 waarnemingen.
- de PO groep, met een gemiddelde naonderzoekperiode van 180 weken, standaard afwijking van 47,2 en 85 gevallen.
- de NO-groep, met een gemiddelde naonderzoekperiode van 316 weken, standaard afwijking van 46,7 en totaal 154 gevallen.

Weber geeft onvoldoende gegevens om het verschil tussen de naonderzoekperiode van W en PO te toetsen, doch het verschil tussen de gemiddelde naonderzoekperiodes is groot genoeg om toch het verwerpen van een  $H_0$  te motiveren.

Ook het verschil tussen de gemiddelde naonderzoekperiodes van de PO en NO groepen is groot. Toetsing met een t-toets levert een  $P < 0,001$  op.

De vergelijkbaarheid van deze steekproeven komt hiermede zodanig in het gedrang, dat aan de consequenties van deze verschillen in naonderzoekperiode een apart hoofdstuk gewijd werd (Hoofdstuk VI).

#### **F Verwaarlozing van de invloed van selectie bij het vergelijken van klinische steekproeven.**

Indien men bij een patiënt de beschikking heeft over twee verschillende behandelings technieken, dan zal men meestal uit deze twee alternatieven een keuze moeten doen. Men kan dat doen op basis van een gevoelsmatige beslissing (pag. 32), die berust op een conglomeraat van informatie waar ervaring en kennis onderdelen van zijn. Men neemt dan aan dat ervaring werd opgedaan in gevallen die voldoende gelijkenis vertoonden met het nu op te lossen probleem, en dat toepassing van deze ervaring verantwoord is.

Met andere woorden de ervaring en de kennis moeten representatief zijn voor het op te lossen keuzeprobleem.

Men kan het beste alternatief ook trachten te berekenen en de keuze doen door middel van een statistische beslissing. In dit geval zal men relevante gegevens bestuderen die men door middel van steekproeven heeft verkregen. Op grond hiervan kan men uitspraken doen over de populaties waaruit de steekproeven op aselechte wijze werden verkregen. Zo kan men een beslissing nemen over de voorkeursbehandeling voor de populaties waaruit de bestudeerde steekproeven werden getrokken. Het is immers een grondeigenschap van de statistiek generaliserende beslissingen te kunnen nemen op basis van incomplete (steekproef) informatie.

Is een dergelijke beslissing nu ook relevant voor het op te lossen individuele keuze-probleem? Indien de patiënt afkomstig kan worden geacht uit een populatie, die identiek is met de populaties van de onderzochte steekproeven, dan is de beslissing juist voor de populatie waaruit hij komt of voor een aselechte steekproef die deze populatie representeert. Eén patiënt is echter maar een heel kleine steekproef en of de beslissing ook voor deze uit één element samengestelde steekproef juist is blijft onzeker.

Dit wil zeggen, dat op welke wijze men ook beslist, de geldigheid van deze beslissing staat of valt met de mate van representativiteit van de beslissingsgronden. De gevoelsmatige beslissing zal vermoedelijk in een aantal gevallen hoog gekwalificeerd zijn met een sterke specificiteit voor het individuele beslissingsprobleem. Zoals reeds op pag. 33 werd uiteengezet is echter de kwaliteit noch het representatief zijn van de verwerkte informatie controleerbaar. De relevantie van de statistische beslissing is sterk afhankelijk van het identiek zijn van de door middel van steekproeven onderzochte populaties en de populatie waarvan de individuele patiënt deel uitmaakt als steekproefelement.

De relevantie van de beslissing voor de individuele patiënt hangt echter af van de representatieve waarde, die aan deze patiënt ten aanzien van de populatie, waaruit hij als steekproef werd 'getrokken', mag worden toegekend. Deze representatieve waarde is het hoogste, indien wordt verondersteld, dat de steekproef van één element is getrokken uit een populatie, bestaande uit nagenoeg identieke elementen. Aangezien de onderzoekpopulatie bij een klinisch onderzoek meestal juist bestaat uit elementen die onderling ongelijk zijn, blijft de geldigheid van de beslissing ten aanzien van de individuele patiënt ook met de statistische methode onzeker. Deze onzekerheid is bij de statistische beslissing echter formuleerbaar en te controleren. De geldigheid van de statistische beslissing ten aanzien van de populatie waaruit de patiënt, beschouwd als steekproefelement, afkomstig is hangt af van het identiek zijn van drie populaties. Namelijk, de twee populaties, die door middel van steekproeven werden onderzocht dienen, behoudens de te beoordelen parameters, identiek te zijn ten opzichte van elkaar en met de populatie waaruit de patiënt afkomstig is.

Indien men zich wil beraden over de geldigheid van de beslissing dan zal men dus moeten nagaan om welke populaties dit gaat.

1. De populatie waaruit de patiënt, als steekproefelement, afkomstig is: Stel men moet een beslissing nemen omtrent de te hanteren behandelingstechniek bij een patiënt met een Letsel van het B.S.G. Men moet kiezen uit een behandeling volgens Weber of een behandeling niet volgens Weber. Men kan deze patiënt dan beschouwen als een steekproefelement uit een populatie bestaande uit personen die allen de eigenschap **L** hebben. Hierbij stelt de eigenschap **L** het hebben van een letsel van het B.S.G. voor. Is genoemde populatie nu identiek met alle personen uit de wereldbevolking die de eigenschap **L** hebben?

Dit is zeker niet het geval. Een ziekenhuis werkt door zijn localisatie en aard selecterend op de wereldbevolking. En dit niet alleen naar ras en klimaat, doch ook naar milieu van de patiënt en type letsel.

Als men nu de selecterende werking van het ziekenhuis waar de patiënt ter behandeling binnenkwam als eigenschap **Z** samenvat, dan kan men de populatie, waar de patiënt als steekproefelement uit afkomstig is, formuleren als die deelverzameling personen uit de wereldbevolking die allen de eigenschappen **Z** en **L** hebben. Tenslotte dient men zich te realiseren, dat de eigenschappen **Z** en **L** niet één enkele vaste parameter voorstellen doch een conglomeraat van factoren, dat voor de behandelingskeuze van **L** relevant wordt geacht. Zo is b.v. de aard van het ongeval dat **L** veroorzaakte een onderdeel van **Z** en de ernst van de weke delen beschadiging een onderdeel van **L**. Dit brengt mee dat de betekenis van **Z** en **L** niet constant is in de tijd. Indien b.v. een nieuw ziekenhuis in de regio wordt geopend, dan kan dat invloed hebben op de selecterende werking, die het eigen ziekenhuis op de populatie van de regio had. Als de verkeersdrukke toeneemt, dan kan het aantal ernstige begeleidende weke delen letsels in de populatie toenemen. Dit wil zeggen, dat de populatie aan veranderingen onderhevig is en dat de patiënt die vandaag binnenkomt, als steekproefelement beschouwd, uit een andere populatie afkomstig is dan de patiënt, die volgend jaar met hetzelfde letsel binnenkomt. Hoe groter dit tijdsinterval is, des te sterker moet men zich afvragen of het nog verantwoord is beide patiënten te beschouwen als steekproefelementen uit dezelfde populatie.

Men zou de populatie waaruit de te behandelen patiënt als steekproefelement getrokken is zijn achtergrondpopulatie kunnen noemen.

2. Ten einde de bovengenoemde beslissing te kunnen nemen doet men een onderzoek. Men verzamelt patiënten, die de eigenschappen **L** en **Z** hebben en past de behandelingstechniek van Weber toe. Deze groep patiënten krijgt dan bovendien de eigenschap **W**.

De patiënten met de eigenschappen **L,Z** en **W** worden gezamenlijk als een steekproef beschouwd. Rekening houdend met de reeds beschreven invloed van de tijd op de achtergrondpopulatie kan men de populatie van deze steekproef formuleren

als de deelverzameling personen uit de wereldbevolking, die in de periode dat de steekproef werd verzameld de eigenschappen **L**, **Z** en **W** hadden. Onderzoekt men deze patiënten na enige tijd en meet men het Resultaat van de behandeling, dan krijgen de patiënten er de eigenschap **R** bij. De achtergrondpopulatie van de steekproef blijft dezelfde, d.w.z. de deelverzameling **die tijdens het verzamelen van de steekproef de eigenschappen L, Z en W had**. De elementen van de populatie hebben de eigenschap **R** (= het geformuleerde behandelingsresultaat) niet, doch het resultaat van de behandeling zou wel gemeten (en daarmee geformuleerd) kunnen worden.

In tweede instantie verzamelt men een andere groep patiënten, die de eigenschappen **L** en **Z** hebben en past een behandeling niet volgens Weber toe ( $=\overline{W}$ ), deze patiënten krijgen er dus de eigenschap  $\overline{W}$  bij. Ook deze patiënten worden als een steekproef beschouwd en wel uit de populatie, die gevormd wordt door die deelverzameling uit de wereldbevolking, die de eigenschappen **L**, **Z** en  $\overline{W}$  hebben, in de periode dat de steekproef wordt verzameld.

Ook de elementen uit deze steekproef worden na enige tijd onderzocht en krijgen er de eigenschap **R** bij. Ook hier blijft hun achtergrondpopulatie dezelfde.

De te beoordelen parameters bij de twee steekproeven zijn **W**,  $\overline{W}$  en **R**. Overigens dienen hun achtergrondpopulaties identiek te zijn. Identiek zijn deze twee populaties alleen indien ook de perioden waarin de twee steekproeven werden verzameld identiek zijn.

Hieruit blijkt, dat steekproeven, die niet in dezelfde periode worden verzameld, nauwelijks met elkaar te vergelijken zijn. D.w.z. voor een goede vergelijking is het nodig dat de patiënten de eigenschappen **L** en **W** of  $\overline{W}$  in dezelfde periode kregen. Immers statistische beslissingen berusten op het berekenen van de kans dat bepaalde verdelingsverschillen kunnen optreden, bij trekking uit identieke populaties. Als nu al bij voorbaat vaststaat, dat de populaties niet identiek zijn, dan gaat veel grond van de statistische beslissing verloren. Vanuit een andere invalshoek werd dit probleem reeds besproken, doch ook nu blijkt de factor tijd een belangrijke rol te spelen.

De twee bij het onderzoek te vergelijken steekproeven waren de groep met de eigenschappen **L**, **Z**, **W** en **R** en de groep met de eigenschappen **L**, **Z**,  $\overline{W}$  en **R**. Hun achtergrondpopulaties staan boven geformuleerd. Daarbij is er echter van uitgegaan dat de groep patiënten, die werd naonderzocht, nog steeds dezelfde groep is als de patiënten, die aanvankelijk werden behandeld.

Dit is echter nooit het geval. Hoe langer het tijdsinterval is tussen oorspronkelijke behandeling en het naonderzoek, des te meer patiënten uit de oorspronkelijke groep kunnen niet voor naonderzoek worden bereikt. De steekproef **L**, **Z** en **W** is dus groter dan de steekproef **L**, **Z**, **W** en **R**, hetzelfde geldt voor de steekproeven met de behandeling  $\overline{W}$ . Dit verschil kan men het **verlies** noemen. Men moet zich nu afvragen of met het optreden van dit verlies ook de achtergrondpopulatie van de steekproef verandert.



Dit behoeft niet noodzakelijkerwijs het geval te zijn. Als dit niet het geval is, wordt door het kleiner worden van de steekproeven alleen het onderscheidend vermogen van het beoordelen van verdelingsverschillen tussen de twee te vergelijken steekproeven geringer. De achtergrondpopulaties veranderen niet indien het verlies aselekt is ten aanzien van die parameters, waarvan men weet of moet aannemen dat ze invloed hebben op de te beoordelen parameters. Men zou dus het verlies op zijn aselectiviteit moeten toetsen ten aanzien van deze parameters. Immers een niet aselekt verlies zou de achtergrondpopulatie van de steekproef op onbekende wijze doen veranderen. Indien men nu toch gaat vergelijken en een uitspraak doet, dan is het niet meer mogelijk om te bepalen of, en voor welke populaties, deze uitspraak geldig is.

Ten einde na te gaan of het verlies aselekt genoemd mag worden, zal men moeten weten hoe het verlies tot stand komt.

Voordat de uiteindelijk te beoordelen steekproef, waarvan de elementen de eigenschappen  $L$ ,  $Z$ ,  $W$  of  $\bar{W}$ , en  $R$  hebben, gevormd is, zijn er talloze momenten geweest, dat verlies kon optreden. De voornaamste hiervan zijn:

- Er kan verschil zijn opgetreden tussen het aantal patiënten dat binnen de periode dat de steekproef zou worden samengesteld, met de eigenschap  $L$  het ziekenhuis binnenkwam, en het aantal patiënten dat als elementen van de steekproef werd geregistreerd.
- In de periode, die verloopt tussen het uitvoeren van de behandelingen  $W$  en  $\bar{W}$  en het moment van het naonderzoek kan een aantal patiënten, die tot de oorspronkelijke steekproef behoorden, overleden zijn.
- In dezelfde periode kan een aantal patiënten voor de te beoordelen steekproef verloren gaan doordat ze door verhuizing niet voor een naonderzoek bereikt kunnen worden.
- Tenslotte kan, vooral ten aanzien van een naonderzoek dat in het ziekenhuis moet plaats vinden, verlies optreden omdat een aantal patiënten ondanks herhaaldelijk verzoek niet voor een naonderzoek verschijnt.

De grootte van het eerstgenoemde verschil hangt nauw samen met de betrouwbaarheid van de registratie van het ziekenhuis, en met de organisatie van het signaleren van patiënten, die relevant zijn voor het onderzoek. Prospectief opgezette onderzoeken zijn hier duidelijk in het voordeel, aangezien de alertheid van de onderzoeker dit verlies belangrijk kan verkleinen, terwijl hij bij een retrospectief onderzoek geheel afhankelijk is van de toegankelijkheid en de volledigheid van de medische archivering. Een typische eigenschap van dit verlies is, dat het niet controleerbaar is. Immers zodra een relevante patiënt wordt gevonden, wordt hij in het onderzoek betrokken en geregistreerd als element van de steekproef, een niet gevonden patiënt blijft onbekend.

Door het bestaan van verschillende vormen van patiëntenregistratie-en archiveringen

is het echter meestal wel mogelijk een schatting te maken van de grootte van het verlies. Zekerheid omtrent compleetheid van de gevormde steekproef is met de gangbare registratiemethoden echter nimmer te krijgen.

Om dezelfde reden is ook de vraag of dit verlies aselect is niet te beantwoorden. De enige manier om zo goed mogelijk tegen een selecterende invloed te waken is zo goed mogelijk voor compleetheid te zorgen. Het is immers zeer goed mogelijk dat b.v. juist de probleempatiënten slecht worden gevonden omdat er vaker over gecorrespondeerd wordt en hun dossiers langer op (of in!) het bureau van de behandelende arts blijven liggen.

Het tweede type verlies, overlijden van patiënten voordat ze zijn naonderzocht, is onvermijdelijk, doch niet aselect. De oudere patiënten hebben nu eenmaal een hogere sterftেকans dan de jongere. Deze selectie speelt geen rol ten opzichte van de vergelijkbaarheid van de te beoordelen steekproeven, doch pas ten aanzien van een patiënt waarbij de onderzoekconclusie moet worden toegepast. De selecterende invloed van het tussentijds overlijden is in beide te vergelijken steekproeven identiek, mits het overlijden los staat van de te beoordelen parameters.

Het derde type verlies, het door verhuizing buiten bereik raken van patiënten is eveneens praktisch onvermijdelijk. Moeilijker is het om een uitspraak over de selecterende werking te doen. Het lijkt niet onmogelijk, dat het sociale milieu of de leeftijd invloed heeft op het verhuisgedrag. Bevolkingsgegevens zouden hier opheldering over kunnen geven. Zolang echter de verhuisreden los staat van de te beoordelen parameters zal de eventuele, maar hypothetische selecterende invloed van dit verlies in beide te vergelijken steekproeven een even sterk effect hebben en dus de vergelijkbaarheid niet beïnvloeden.

Het vierde type verlies, het niet verschijnen van patiënten ondanks herhaaldelijk verzoek om zich te laten naonderzoeken, is het lastigste verlies. Allerlei argumenten zijn aan te voeren om er een selecterende werking aan toe te kennen. Zo is er redelijkerwijze een verschil te verwachten tussen het type patiënten dat direct na de eerste oproep verschijnt en de patiënten, die pas na herhaaldelijk oproepen, of telefonisch contact, bereid zijn te komen. Over deze verschillen is weinig bekend zodat het niet verantwoord is om de hypothese te hanteren dat de argumenten, die de patiënten hebben om niet voor een onderzoek te verschijnen, geen verband houden met de in het onderzoek te beoordelen parameters. M.a.w. dat het verlies aselect genoemd mag worden.

Ten aanzien van twee te vergelijken groepen is het dan ook van belang om dit verlies in beide steekproeven op dezelfde wijze te benaderen om het selecterend effect zoveel mogelijk identiek te doen zijn. Dit kan gebeuren door steeds op standaardwijze op te roepen, b.v. vier keer schriftelijk en één keer telefonisch. Van het dan overblijvende verlies kan men zich door middel van een steekproef een oordeel vormen. De patiënten uit deze steekproef zouden dan door middel van huisbezoek onderzocht moeten worden.

Uiteindelijk moet men dus tot de conclusie komen, dat het verlies weliswaar zo klein mogelijk kan worden gehouden, doch dat het nooit geheel te vermijden is, terwijl het bovendien niet aselekt is. Zolang het selecterend effect alleen parameters geldt, die los staan van de bij een onderzoek te beoordelen parameters, mag het ten aanzien van de te vergelijken steekproeven als identiek worden beschouwd en komt de vergelijkbaarheid er niet door in gevaar.

Aangezien men aan het verlies een selecterende werking moet toekennen moet men deze selectie in de eigenschappen van de elementen van de te vergelijken steekproeven tot uitdrukking laten komen. Indien men de selecterende invloed van het verlies in de eerste steekproef met  $V_1$  samenvat en die in de tweede steekproef met  $V_2$  dan zijn de eigenschappen van de genoemde elementen:

$L, Z, W, R$  en  $V_1$  en  $L, Z, \bar{W}, R$  en  $V_2$ . Om statistische vergelijking verantwoord te doen zijn dienen  $V_1$  en  $V_2$  identiek te zijn.

Stel deze identieke selecterende werking van het verlies op  $V$ .

Indien men nu op basis van de statistische vergelijking een conclusie trekt over de behandelingstechnieken  $W$  en  $\bar{W}$ , dan mag deze conclusie generaliserend worden uitgebreid over de populatie, die wordt gevormd door die deelverzameling uit de wereldbevolking, waarvan de elementen de eigenschappen  $L$  en  $Z$  hadden in de periode dat de oorspronkelijke steekproef werd samengesteld en nu bovendien de eigenschap  $V$  hebben gekregen.

Als we nu terug keren tot de oorspronkelijk te behandelen patiënt van pag. ... dan bleek hij, als steekproefelement een populatie te vertegenwoordigen, die uit elementen bestond met de eigenschappen  $L$  en  $Z$ . Deze populatie is niet identiek met de door middel van steekproeven onderzochte populatie. Strikt genomen is de onderzoekconclusie dus niet geldig voor zijn achtergrondpopulatie en daarmee voor de patiënt. Indien men nu toch de onderzoekconclusie op de patiënt van toepassing brengt dan verwaarloost men:

- Het effect op  $L$  en  $Z$  van het tijdsverschil tussen het vormen van de onderzochte steekproeven en het binnenkomen van de te behandelen patiënt.
- Het effect van  $V$  op de relevantie van de onderzoekconclusie ten aanzien van de te behandelen patiënt.

In veel gevallen zal het verantwoord zijn om deze invloeden te verwaarlozen, men dient zich er echter wel van bewust te blijven. Het effect van het tijdsverschil zal waarschijnlijk met de tijd toenemen en zo kan het ter beschikking komen van nieuwe behandelingsmogelijkheden de eigenschap  $Z$  zo sterk veranderen, dat relevant achten van de onderzoekconclusie niet meer verantwoord is.

Ook het verwaarlozen van  $V$  kan een gevaar inhouden. Zelfs indien bij het ideale onderzoek alleen de selecterende invloed van het verlies door overlijden van patiënten is overgebleven moet men zich nog afvragen of de oudere patiënt, die met de eigenschap  $L$  het ziekenhuis binnenkomt wel voldoende in de onderzoekconclusie is vertegenwoordigd.

Tenslotte blijft de vraag over naar de geldigheid van de conclusie voor de uit één element bestaande steekproef (de te behandelen patiënt). Indien van de patiënt niet meer informatie beschikbaar is dan dat hij de eigenschappen **L** en **Z** heeft en hij dus tot de relevante achtergrondpopulatie behoort, is de onderzoekconclusie ook voor de individuele patiënt zonder meer geldig. **L** en **Z** vertegenwoordigen echter conglomeraten van factoren, die invloed hebben op het beoordeelde behandelingsresultaat. Een globale beoordeling van het gehele conglomeraat geeft geen informatie over de individuele factoren, zodat het mogelijk is dat deze zich in bijzondere gevallen bij een patiënt ongemerkt onderling zodanig opstellen dat de onderzoekconclusie niet de beste genezingskans waarborgt. Zolang over individuele factoren geen gedetailleerde kennis beschikbaar is, kan slechts gevoelsmatig beslist worden of een patiënt tot deze bijzondere categorie behoort.

In dit verband is het van belang, dat zoveel mogelijk kennis wordt verzameld over alle factoren, die van invloed zijn bij het beoordelen van te vergelijken steekproeven. Hiermee zou het mogelijk zijn deelsteekproeven samen te stellen, die uit elementen bestaan, die onderling meer gelijk zijn. Conclusies uit onderzoek van dergelijke steekproeven getrokken zijn geldig voor populaties, die op hun beurt uit onderling ongeveer gelijke elementen bestaan. Deze conclusies zijn dan met groter waarschijnlijkheid geldig voor iedere individuele patiënt.

Laten we nog eens terugkeren naar de steekproeven, die in het kader van het onderzoek met elkaar werden vergeleken. De eigenschappen van hun elementen, **L** en **Z**, moesten identiek zijn in het kader van hun vergelijkbaarheid. De eigenschap **R** was het te beoordelen resultaat van de behandeling van **L**. Indien een significant verdeelingsverschil van de eigenschap **R** tussen beide steekproeven werd gevonden, dan is de bij de vergelijking van de twee steekproeven op te stellen  $H_0$  te verwerpen ten gunste van  $H_1$  (Zie hoofdstuk II) hetgeen wil zeggen, dat beslist wordt, dat de twee steekproeven niet uit identieke populaties afkomstig zijn. Het niet identiek zijn van de twee achtergrondpopulaties wordt dan toegeschreven aan het enige bekende verschil tussen de twee populaties, namelijk **W** en  $\bar{W}$ . Deze interpretatie lijkt gerechtvaardigd omdat aannemelijk lijkt dat **L**, **Z** en **W** of  $\bar{W}$  in combinatie met elkaar **R** bepalen. Indien men zich nu herinnert, dat **L** de te behandelen afwijking voorstelt, **Z** alle relevante factoren, die aan het ziekenhuis gebonden zijn en **W** en  $\bar{W}$  de twee met elkaar te vergelijken behandelingsalternatieven, dan blijkt direct, dat de boven gestelde interpretatie niet juist is. Immers het resultaat van een afwijking is nooit alleen afhankelijk van de afwijking zelf en de behandeling. De patiënt zelf herbergt altijd een groot aantal factoren, die weliswaar niets met de afwijking te maken hebben, doch wel mede van invloed zijn op zijn genezingskans. Zo zullen leeftijd en werkkring mede bepalen of patiënt last houdt van een enkelfractuur en ook zijn kans op complicaties beïnvloeden.

Als we dit aan de patiënten gebonden, maar los van **L** staande conglomeraat van

factoren samenvatten als **P**, dan blijken de twee te vergelijken steekproeven identiek te moeten zijn ten aanzien van **L**, **Z** en **P**. Indien de steekproeven beiden op strikt aselechte wijze tot stand zijn gekomen, dan behoeft men zich hierover geen zorgen te maken. Wij zagen echter reeds, dat in beide steekproeven een selecterend verlies kan optreden, genoemd resp.  $V_1$  en  $V_2$ , dat het identiek zijn van **L**, **Z** en **P** bedreigt.

Of aan de voorwaarde voor vergelijkbaarheid van de steekproeven ( $V_1 = V_2$ ) wordt voldaan is niet direct te beoordelen. Om hierover een indruk te krijgen kan het verschil betreffende **L**, **Z** en **P** wat door het effect van  $V_1$  en  $V_2$  zou kunnen zijn ontstaan, beoordeeld worden. Hiertoe moeten de factoren, die **R** beïnvloeden en respectievelijk door **L**, **Z** en **P** werden voorgesteld in beide steekproeven worden gemeten, waarna een eventueel verdelingsverschil tussen de beide steekproeven kan worden getoetst. Wordt ten aanzien van enkele van deze factoren een significant verschil gevonden dan moet men aannemen, dat  $V_1$  en  $V_2$  althans ten aanzien van deze factoren niet gelijk waren. Dit behoeft niet altijd te leiden tot volledige onvergelijkbaarheid van de twee steekproeven, doch het leidt wel tot een aantal interpretatierestricties bij het trekken van conclusies uit de vergelijking.

Indien men twee steekproeven vergelijkt en uit deze vergelijking conclusies trekt, zonder de boven beschreven toetsing uit te voeren ten aanzien van het al of niet gelijk zijn van  $V_1$  en  $V_2$  dan hanteert men stilzwijgend de hypothese dat eventuele verschillen tussen de steekproeven wat betreft **L**, **Z** en **P** verwaarloosbaar klein zijn. Met andere woorden men neemt aan dat alle besproken selecterende invloeden een onbelangrijk effect hebben gehad ten opzichte van de getrokken conclusie.

Vergelijkt men op deze wijze twee steekproeven, die uit verschillende ziekenhuizen afkomstig zijn dan is het zeer de vraag of de veronderstelling dat **Z** in beide steekproeven identiek is, verantwoord is. Wordt in de twee te vergelijken steekproeven een verschillende klassificatie van de te beoordelen afwijking gebruikt, dan weet men zeker dat **L** bij de steekproeven niet identiek is. Bestaat er bovendien nog een tijdsinterval tussen de periodes dat de steekproeven werden samengesteld, dan wordt nog een selecterende invloed verwaarloosd. Een conclusie die op basis van een dergelijke vergelijking tot stand komt staat op zeer losse gronden, is hypothetisch en kan totaal onjuist zijn uitsluitend door verwaarlozing van allerlei selecterende invloeden.

Als Weber de resultaten van zijn behandelingstechniek, samengevat in zijn eindscores, vergelijkt met gegevens uit de literatuur, dan wordt op de zojuist beschreven wijze, de selecterende invloed op de vergeleken steekproeven verwaarloosd. De conclusie, die hij uit deze vergelijking trekt is dus alleen al op deze gronden aanvechtbaar.

De beschouwingen over de verwaarlozing van selectie hebben bij het beoordelen van het eigen onderzoek de volgende consequenties:

- De vergelijkbaarheid van de steekproeven dient te worden beoordeeld.

- Hiertoe zullen een aantal relevante factoren op hun verdelingen binnen de steekproeven worden vergeleken.
- Indien tot een selecterend effect ten aanzien van deze factoren besloten moet worden, dan zullen interpretatierestricties worden geformuleerd, die in de eindconclusies moeten worden verwerkt.
- Het verlies (V) zal in beide te vergelijken steekproeven op zijn selecterende invloed worden getoetst.

Aan het slot van dit hoofdstuk wil ik nog eens (zie de Inleiding) wijzen op het tweede deel van de literatuurlijst. Hierin is literatuur verzameld, die voor nadere oriëntatie over de in dit hoofdstuk besproken problemen kan dienen.

Aanhaling van auteurs in dit hoofdstuk achtte ik niet op zijn plaats, aangezien alle redeneringen een fundamentele logica als basis hebben en geënt zijn op de grondbeginselen van de beslissingstheorie.

De aanhaling van enkele auteurs op dit gebied zou alle andere tekort hebben gedaan.

Hoewel over methodologische problemen bij de klinische research al heel wat is gepubliceerd (Cruchet, Martini, Trotter, Mainland, Croxton, De Jonge, Bradford Hill, Wibaut e.v.a.) is een beschouwing, zoals in het voorafgaande hoofdstuk, mij niet bekend. De meeste schrijvers zoeken de problemen in hoofdzaak in principiële fouten in de bewerking van de verzamelde gegevens. Aan de gevolgen van onnauwkeurigheden van de inhoudelijke betekenis van de onderzoekgegevens wordt te zelden meer dan globale aandacht besteed. Wel waarschuwt Mainland (1960) tegen het overmatig vertrouwen in statistische toetsingen, toegepast op eindscores. Deze geven, stelt hij, niet zelden een exactheid, die door de gebruikte onderzoeksmethode niet voldoende wordt ondersteund. Zo ontstaat slechts een schijnexactheid, die beter achterwege kan blijven.

Vaak wijt men het falen van mathematische onderzoekmodellen bij het oplossen van klinische problemen aan een principiële onmogelijkheid om deze problemen goed in een model onder te brengen. De ervaring is echter nog maar klein en voorlopig lijkt het verstandiger om aan te nemen, dat het falen niet aan principiële onmogelijkheid, doch aan een incompleet en/of onjuist opgebouwd model moet worden toegeschreven.

## Naonderzoek formulier

Serie PO=1 NO=2 .....

Nummer .....

Geslacht

vrouw=0 man=1 .....

Leeftijd .....

Beroep Licht=0 Matig=1 Zwaar=2 .....

Vrije tijdsbelasting Licht=0 Matig=1 Zwaar=2 .....

Jaar en dagnummer ongeval .....

Kant

links Neen=0 Ja=1 .....

rechts Neen=0 Ja=1 .....

Complicatie .....

Primaire dislocatie .....

### Type fractuur

Indeling mogelijk

0=alleen vlg. Weber

Weber

1=alleen vlg. L.H. 2=niet indeelbaar

Luxatiefractuur type A Neen=0 Ja=1 .....

id. type B Neen=0 Ja=1 .....

id. type C Neen=0 Ja=1 .....

Stauchungsfractuur type A Neen=0 Ja=1 .....

id. type B Neen=0 Ja=1 .....

id. type C Neen=0 Ja=1 .....

Andere enkelfractuur type A Neen=0 Ja=1 .....

id. type B Neen=0 Ja=1 .....

id. type C Neen=0 Ja=1 .....

### Lauge Hansen

SE fractuur type I Neen=0 Ja=1 Ongeveer=2 .....

id. type II Neen=0 Ja=1 Ongeveer=2 .....

id. type III Neen=0 Ja=1 Ongeveer=2 .....

id. type IV Neen=0 Ja=1 Ongeveer=2 .....

PE fractuur type I Neen=0 Ja=1 Ongeveer=2 .....

id. type II Neen=0 Ja=1 Ongeveer=2 .....

id. type III Neen=0 Ja=1 Ongeveer=2 .....

id. type IV Neen=0 Ja=1 Ongeveer=2 .....

SA fractuur type I Neen=0 Ja=1 Ongeveer=2 .....

id. type II Neen=0 Ja=1 Ongeveer=2 .....

PA fractuur type I Neen=0 Ja=1 Ongeveer=2 .....

id. type II Neen=0 Ja=1 Ongeveer=2 .....

id. type III Neen=0 Ja=1 Ongeveer=2 .....

PD fractuur type I Neen=0 Ja=1 Ongeveer=2 .....

id. type II Neen=0 Ja=1 Ongeveer=2 .....

id. type III Neen=0 Ja=1 Ongeveer=2 .....

id. type IV Neen=0 Ja=1 Ongeveer=2 .....

### Earle

Neen=0 gering=1  $\perp$  1/5=2  $\perp$  1/3=3 .....

### Talus

Neen=0 gering=1 ernstig=2 .....

## Communitie

mediaal Neen=0 gering=1 ernstig=2 .....  
lateraal Neen=0 gering=1 ernstig=2 .....

## Bandletsels

voorste syndesmoose Neen=0 Ja=1 .....  
achterste syndesmoose Neen=0 Ja=1 .....  
mediaal Neen=0 Partieel=1 Totaal=2 .....  
lateraal Neen=0 1/3=1 2/3=2 Totaal=3 .....

## Operatie Neen=0 Ja=1 .....

volgens Weber Neen=0 Ja=1 .....  
Direct=1 Binnen 8 uur=2 Later=3 .....  
Interval ongeval-operatie .....  
Operateur Ik=0 ander=1 .....  
Bloedleegte aantal kwartieren ( $\lambda$  n.v.t.) .....  
Moeilijkheid geen probleem=0 matig=1 moeilijk=2 ( $\lambda$  n.v.t.) .....  
Suprasyndesmale fixatie Neen=0 Ja=1 .....

## P.O. genezing p.p. Ja=0 Neen=1 .....

Sec. ingreep nodig Neen=0 Ja=1 .....  
Complicaties Neen=0 Ja=1 .....  
longembolie =1 wondinfectie =2 .....  
osteomyelitis =3 thrombose =4 .....  
p.o. overleden =5 andere =6 .....

## Direct resultaat

Anatomie perfect=0 goed=1 voldoende=2 slecht=3 .....  
Luxatiestand  
Lateraal In mm .....  
mediaal In mm .....  
voren In mm .....  
achteren In mm .....  
kanteling varus aantal graden .....  
valgus aantal graden .....  
vork te wijd Neen=0 Matig=1 Ernstig=2 .....  
Stabiliteit Oefenstabil=0 Niet oefenstabil=1 .....  
Niveau verschil tertius In mm .....

## Morbiditeit

Werken in% na weken .....  
Werken 100% na weken .....  
Lopen zonder klachten Neen=0 Ja=1 .....  
Tijd in weken .....

## Loopklachten bij eind behandeling

Geen=0 Gering=1 Matig=2 Ernstig=3 .....  
Eind behandeling in weken .....

## Materiaal verwijderd aantal weken $\lambda$ n.v.t. ....

onder lokaal anaesthesie Ja=0 Neen=1 .....  
onder algemene anaesthesie Neen=0 Ja=1 .....  
wondgenezing per primam Ja=0 Neen=1 .....  
secundaire ingreep nodig Neen=0 Ja=1 .....



tijd in weken .....

pijn .....	
lopen .....	
activiteit .....	
anatomie .....	
bov.spronggewricht .....	
ond.spronggewricht .....	
<b>totaal .....</b>	

Resultaat      Zeer goed=0      Goed=1      Slecht=2      .....

Subjectief  
 patiënt tevreden=0      Matig tevreden=1      ontevreden=2 .....  
 totaal subjectief (PP 1, 2 en 3 van Weber en "tevredenheid") .....

para-achillairrelief fraai=0 verstreken=1 verbreding=2 bol=3 .....  
 symmetrie homolat.het beste=0 li=re=1 homolat. het slechtst=2 .....

lengtegewelf	normaal=0	vlakker=1	vollig verdwenen=2	valgus=3	.....
holvoet	Neen=0	Ja=1	.....	.....	.....
voorvoet	normaal=0	spreiervoet=1	hamerstand	tenen=2	.....
superductusstand	tenen	Neen=0	Ja=1	.....	.....
symmetrie	homolat. het beste=0	homolat. het slechtst=2	li=re=1	.....	.....

voet heterolateraal	Neen=0	Ja=1	.....
homolateraal	Neen=0	Ja=1	.....
been heterolateraal	Neen=0	Ja=1	.....
homolateraal	Neen=0	Ja=1	.....

abnormale vaattekening				
homolateraal	Neen=0	Matig=1	Sterk=2	.....
heterolateraal	Neen=0	Matig=1	Sterk=2	.....

homolateraal	Neen=0	Matig=1	Sterk=2	.....
heterolateraal	Neen=0	Matig=1	Sterk=2	.....

Neen=0   Matiq=1   Sterk=2   .....

## homolat./heterolat. ....

0=neen    1=ja, onpijnlijk    2=ja, pijnlijk    .....

"slecht weer" klachten      Neen=0    Ja=1    .....

"weefsel" prae-achillair      Neen=0      Weinig=1      Veel=2

## "Drukpijn"

mediale band	Neen=0	Ja=1	.....
laterale band	Neen=0	Ja=1	.....
voorste syndesmoose	Neen=0	Ja=1	.....
achterste syndesmoose	Neen=0	Ja=1	.....
mediale malleolus	Neen=0	Ja=1	.....
laterale malleolus	Neen=0	Ja=1	.....

## Röntgenbevindingen

Restdislocaties			
laterale malleolus		In mm.....	
mediale malleolus		In mm.....	
mediale tibiahoek		In mm.....	
malleolus tertius niveau verschil		In mm.....	
spleet in gewrichtsvlak		In mm.....	

## Restluxatiestand

lateraal	Neen=0	etc. in mm	.....
mediaal in mm			.....
voren in mm			.....
achteren in mm			.....
kanteling varus in graden			.....
valgus in graden			.....
vork te wijd	Neen=0	Matig=1	Ernstig=2 .....

## Arthrose

talo-tibiale sclerose			
mediaal	Neen=0	Matig=1	Ernstig=2 .....
lateraal	Neen=0	Matig=1	Ernstig=2 .....
vernauwing gewrichtsspleet			
talo-tibiaal			
mediaal		In mm.....	
lateraal		In mm.....	
voorkant		In mm.....	
achterkant		In mm.....	
lateraal		In mm.....	
mediaal		In mm.....	
tibio-fibulair		In mm.....	

## Onregelmatigheid van het gewricht

talo-tibiaal			
mediaal	Neen=0	Matig=1	Ernstig=2 .....
lateraal	Neen=0	Matig=1	Ernstig=2 .....
voorkant	Neen=0	Matig=1	Ernstig=2 .....
achterkant	Neen=0	Matig=1	Ernstig=2 .....
mediaal	Neen=0	Matig=1	Ernstig=2 .....
lateraal	Neen=0	Matig=1	Ernstig=2 .....
tibio-fibulair	Neen=0	Matig=1	Ernstig=2 .....

## Osteophytvorming

tibia voorkant	Neen=0	Matig=1	Ernstig=2 .....
achterkant	Neen=0	Matig=1	Ernstig=2 .....
malleolus med.	Neen=0	Matig=1	Ernstig=2 .....
malleolus lat.	Neen=0	Matig=1	Ernstig=2 .....

**Botstructuur**

tibia	gaaf=0	desorganisatie matig=1	idem ernstig=2	.....
	"krachtlijnen"	Neen=0	Ja=1	.....
malleolus lat.	gaaf=0	desorganisatie matig=1	idem ernstig=2	.....
	"krachtlijnen"	Neen=0	Ja=1	.....
malleolus med.	gaaf=0	desorganisatie matig=1	idem ernstig=2	.....
	"krachtlijnen"	Neen=0	Ja=1	.....

**Verkalking in de banden**

mediaal	Neen=0	Ja=1	.....
lateraal	Neen=0	Ja=1	.....
voor	Neen=0	Ja=1	.....
achter	Neen=0	Ja=1	.....

**Verbening t.p.v. de suprasyndesmale Schroef**

$\Delta$  = n.v.t.    0=neen    1=verbening    2=synostose    .....

**Stabiliteit banden**

mediaal (in graden stresshoek $\Delta$ = n.v.t.)	.....
lateraal (in graden stresshoek $\Delta$ = n.v.t.)	.....
vork te breed	Neen=0    Ja=1    .....

**Voor-achterwaartse instabiliteit**

Neen=0	Ja=1	.....
--------	------	-------

## Hoofdstuk IV

### DE GEGEVENS

De bijlage geeft een volledige lijst met de per patiënt verzamelde gegevens.

Wat hun vorm van registratie en kwantificatie betreft, zijn ze grofweg te onderscheiden in 'Ja-Nee' gegevens, waarbij alleen vastgelegd wordt, dat het betrokken fenomeen al of niet afwezig is en de mate van aanwezigheid niet nader wordt vastgelegd, gegevens waarbij door een rangvolgorde (geordend categorisch systeem) een kwantificatie wordt nagestreefd, gegevens, die middels een intervalschaal konden worden gekwantificeerd, coderingen (nominaal categorisch systeem), waarbij een verschillende getalwaarde van de variabele een verschillende begripsinhoud heeft zonder kwantitatieve betekenis en somscores, waarbij de getalwaarden van een aantal variabelen, die geacht worden min of meer bij elkaar te behoren, per patiënt bij elkaar worden geteld en zo de waarde van de somscore bepalen.

#### *'Ja-Nee' gegevens:*

De variabelen: 3, 9, 10, 12 t/m 37, 42, 43, 46, 47, 54, 55, 57, 69, 70, 78, 83, 84, 85, 86, 101, 103, 105, 106, 107, 108, 116, 117, 120, 121, 122, 123, 124, 125, 159, 161, 163, 164, 165, 166, 167, 173.

#### *Rangvolgorde:*

De variabelen: 5, 6, 38, 39, 40, 41, 44, 45, 48, 53, 59, 67, 80, 96, 98, 99, 100, 102, 104, 109, 110, 111, 112, 113, 114, 115, 118, 119, 138, 139, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 160, 162, 168.

#### *Intervalschaal:*

De variabelen: 4, 7, 8, 49, 50, 52, 60, 61, 62, 63, 64, 65, 66, 68, 71 t/m 77, 79, 81, 82, 87, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 140, 141, 142, 143, 144, 145, 146, 169, 170, 171, 172, 174.

### *Coderingen:*

De variabelen: 1, 2, 11, 51, 58.

### *Beoordelingsschema van Weber:*

De variabelen: 88, 89, 90, 91, 92, 93, 95.

### *Somscores:*

De variabelen: 94 (Weber), 97 (subjectief), 175 (objectief).

Rangvolgordes zijn, zoals in het vorige hoofdstuk al werd besproken, nogal eens aanvechtbaar en meestal arbitrair. Ze verdienen daarom een aparte bespreking. De betrouwbaarheid van de indelingen kan blijken uit de later te bespreken relaties.

De beroeps- en vrijetijdsbelasting van het B.S.G. werd gedurende het naonderzoek interview als licht, matig of zwaar beoordeeld. Door de grote gevarieerdheid van de beroepen, waardoor het vaak nauwelijks mogelijk is een duidelijk inzicht te krijgen in de aard van het werk, was het onmogelijk een strakke beoordeling te verkrijgen en werd meestal ad hoc beslist. Toch zijn er wel enkele vaste uitspraken mogelijk; zo werd grondwerk als zware beroepsbelasting en voetbal als zware vrije tijdsbelasting geklassificeerd (Variabelen 5 & 6).

De volgorde bij het tertiusfragment (Var. 38) is eenvoudig. Een klein fragment kreeg de waarde 1, minstens  $\frac{1}{5}$  van de sagittale gewrichtsdoorsnede kreeg een 2, terwijl minstens  $\frac{1}{3}$  een 3 kreeg.

Het talusletsel (39) is moeilijker te meten en bovendien gevarieerder van vorm; het werd ad hoc als gering of ernstig beoordeeld. (Var. 40 en 41).

De verbrijzeling aan mediale en laterale zijde werd steeds aan de hand van de ongevalsfoto als gering, matig of ernstig beoordeeld.

De bandletsels (Var. 44. en 45) werden meestal aan de hand van de operatiegegevens als partieel of totaal (mediaal) vastgelegd, waarbij aan de laterale kant een onderscheid werd gemaakt tussen  $\frac{1}{3}$ ,  $\frac{2}{3}$  en totaal. Hierbij kan toch niet van een intervalschaal gesproken worden, omdat de drie delen van de laterale band onderling niet equivalent zijn.

Variabele 48 geeft aan of direct na binnenkomst werd geopereerd (0), ofwel binnen 8 uur (1), of later (2).

De moeilijkheid van de operatie werd aan de hand van het operatieverslag beoordeeld (var. 53). Indien het niet zonder meer duidelijk was of en hoe een stabiele osteosynthese kon worden verkregen, werd de operatie matig moeilijk beschouwd; indien de eerste poging mislukte en de osteosynthese, c.q. de repositie over moest, dan werd van een moeilijke operatie gesproken.

Het anatomisch resultaat (var. 59) werd goed gevonden (1) indien geen enkele dislocatie groter was 1 mm; slecht indien dislocaties bestonden  $> 3$  mm (3), rest-dislocaties daartussenin kregen een 2.

De primaire dislocatie, voor zover te beoordelen op de ongevalsfoto, volgde de volgende schaal, waarbij de grootste afwijking de codering bepaalt (var.67):

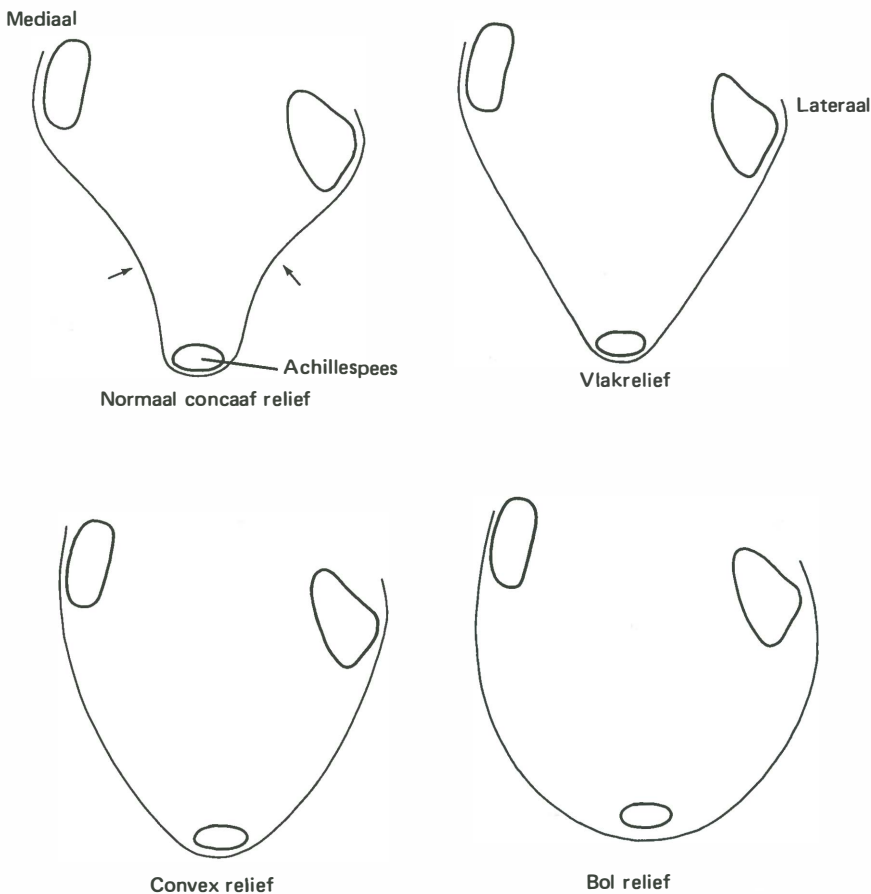
- 0 = geen dislocatie
- 1 = geen (sub) luxaties
  - tot 2 mm verkorting van de fibula
  - tot 2 mm spleetjes van een Ausrisz-fractuur
  - tot 1 mm diastase van het 'tertius' fragment.
  - geen niveauverschil bij een tertiusfragment
  - geen vorkverwijding
  - tot 2 mm anteropositie van een malleolus medialis fragment
- 2 = lateralisatie tot 3 mm
  - kanteling van de talus tot  $3^{\circ}$
  - dislocaties tot 4 mm
  - niveauverschil van het tertius fragment tot 2 mm
  - vorkverwijding tot 3 mm
  - geen luxatie
- 3 = dislocaties  $> 4$  mm
  - subluxaties  $\leq 1/2$
- 4 = geen luxaties
  - subluxatie  $> 1/2$
- 5 = totale luxatie

Bij het naonderzoek werden de klachten bij het lopen in aansluiting aan de behandeling (var. 80) geverifieerd tijdens het interview; naar aanleiding hiervan en naar aanleiding van de notities in de dossiers werden zij ad hoc als gering, matig of ernstig beoordeeld.

Variabele 96, de subjectieve tevredenheid van de patiënt, werd alleen met een nul gehonoreerd, indien volmondig bevestigd werd dat hij/zij volledig met het bereikte resultaat tevreden was. Werd deze bevestiging met enige terughoudendheid gegeven,

dan werd matige tevredenheid genoteerd, terwijl 'wel tevreden maar ik heb nog wel last.....' en minder tevreden uitingen als ontevreden werden geregistreerd.

Het para-achillaire relief, (Variabele 98) is meestal concaaf. Indien er een duidelijk concaaf relief was waar te nemen en dit bovendien niet verstreken was ten opzichte van de gezonde kant, werd een nul genoteerd. Minder concaaf dan de gezonde zijde of een vlakke begrenzing van het gebied tussen achillespees malleoli kreeg een één, een convexe begrenzing een twee, terwijl een drie werd geregistreerd indien dit gebied bol genoemd kon worden. (zie fig. 5)



Figuur 5. 'Para-achillair relief'. Horizontale doorsnede ter hoogte van de malleoli.

Voor de wijze, waarop de symmetrie van het para-achillaire relief (var. 99) werd genoteerd, raadplege men het naonderzoekformulier.

Ook de wijze, waarop variabelen 100, 102 en 104 zijn genoteerd, moge blijken uit het naonderzoekformulier.

De variabelen 109 t/m 113, betreffende de pathologische vaattekeningen en pigmentaties, werden alleen indien juist aanwezig met 'matig' gekwantificeerd. Indien direct waarneembaar werden ze met 'sterk' (2) aangegeven.

De functie van het voorste spronggewricht, het gewricht van Chopart (var. 114), werd aangegeven in het verschil tussen gezonde en zieke zijde. Bij nul bestond er geen verschil, één geeft aan, dat naar schatting 1/10 van de bewegingsmogelijkheid van de gezonde zijde, aan de gefractureerde kant ontbreekt. Bij een twee ontbreekt 2/10 etc., zodat bij een negen een nog maar juist waarneembare bewegingsmogelijkheid bestaat

Variabele 115 geeft met een één de aanwezigheid van een cysteuze zwelling aan over het gebied van de voorste syndesmose. Het viel reeds voor de aanvang van het formele onderzoek op, dat patiënten, die een enkelfractuur hadden gehad, nogal eens een soms pijnlijke zwelling vertoonden over de voorste syndesmose. Deze zwelling fluctueert meestal, soms voelt hij in hoofdzaak pasteus aan. Dikwijls is een knispering waar te nemen. Dit symptoom werd bij het naonderzoek als 'cysteuze zwelling' onderscheiden.

Variabele 118, de bekende startpijn, werd tijdens het interview ad hoc als 'even', 'matig' of 'sterk' vastgelegd, dan wel als afwezig.

Tussen de achterzijde van de distale tibia en de voorzijde van de achillespees bevindt zich meestal zo weinig massa, dat de mediale en laterale huid voor de achillespees met de palperende duim en wijsvinger tegen elkaar is te brengen. Zodanig, dat de indruk bestaat dat er nagenoeg geen weefsel tussen de palperende vingers aanwezig is. Indien dit zo was en indien er ook geen verschil was tussen links en rechts, dan werd een nul geregistreerd voor variabele 119. De beslissing tussen 'matig' of 'veel' werd ad hoc genomen.

Variabelen 138 en 139, alsmede 147 t/m 157 geven nauwelijks objectief te meten waarnemingen aan de röntgenfoto aan. De beslissing tussen matig en sterk werd steeds ad hoc genomen.

Hetzelfde geldt voor variabelen 158, 160 en 162.



Variabele 168 geeft de calcificatie aan ter hoogte van een eventueel gebruikte suprasyndesmale fixatie (schroef). Indien een complete synostose bestond, werd een twee genoteerd.

## Hoofdstuk V

### VERGELIJKBAARHEID, REPRESENTATIVITEIT, TOETSINGSKRITERIA

Zoals reeds in hoofdstuk III werd besproken dient de eigenlijke bewerking van de onderzoeksgegevens te worden voorafgegaan door onderzoek naar de vergelijkbaarheid van de groepen. Vervolgens dient besproken te worden in hoeverre het onderzochte gegevensmateriaal representatief genoemd mag worden, waarna tenslotte de toetsingscriteria een bespreking behoeven.

#### A *Vergelijkbaarheid*

Er werden 97 patiënten met een luxatiefractuur van de enkel volgens de techniek van Weber geopereerd. Van deze 97 patiënten konden er 85 compleet naonderzocht worden (PO groep).

Uit de jaren 1962, '63 en '64 konden 235 patiënten, die destijds wegens een enkelfractuur in het ziekenhuis waren behandeld (meestal conservatief, een enkele maal operatief) uit het archief worden verzameld. Hiervan konden tijdens het onderzoek 156 patiënten worden naonderzocht (NO groep):

Tabel 3. Samenstelling en oorsprong van de onderzochte groepen patiënten.

	PO groep	NO groep
Beschikbaar	97	235
Overleden	5	21
Opgeroepen	92	214
Naonderzocht	85	156
Vertrokken	7	27
	92	183

Bij de NO groep bleven er dus 31 patiënten over, die ondanks herhaalde oproep niet op de uitnodiging reageerden of zeiden niet te kunnen komen. Oorspronkelijk waren dat 41 patiënten, doch door middel van huisbezoek werd een steekproef van 10 gevallen uit deze groep getrokken. De steekproef wordt bij het volgend onderdeel van dit hoofdstuk besproken.

Zoals in hoofdstuk III werd besproken, dienen relevante gegevens uit de PO en NO groep met elkaar vergeleken te worden wat betreft hun verdelingen binnen de twee groepen. Dit om een indruk te krijgen over de rol die de selecterende invloed van het verlies speelt bij de vergelijkbaarheid van de PO en NO groepen.

Bij deze vergelijking bleken de verdelingen van de volgende variabelen significant te verschillen:

zwaarte beroepen	(var. 5)
zwaarte vrijetijd belasting	(var. 6)
verhouding tussen het aantal A-, B- en C-fracturen	(var. 12, 13 en 14)
letsels van de malleolus tertius	(var. 38)
primaire dislocatie	(var. 67)
naonderzoekperiode (in weken)	(var. 87)

De exacte frequenties (F) van deze variabelen waren bij de twee te vergelijken steekproeven (PO en NO):

Tabel 4. Exacte aantallen (frequenties) van bij de PO en NO groepen verschillend verdeelde variabelen.

variabele nr.	F <sub>po</sub>	F <sub>no</sub>
5	34 (0)	71 (0)
	42 (1)	69 (1)
	19 (2)	14 (2)
6	49 (0)	108 (0)
	45 (1)	48 (1)
	1 (2)	— (2)
12, 13 en 14	8, 46 en 33	23, 74 en 21
38	42 (0)	94 (0)
	38 (1)	33 (1)
	5 (2)	16 (2)
	9 (3)	12 (3)
67	18 (1)	32 (0)
	24 (2)	31 (1)
	28 (3)	45 (2)
	8 (4)	24 (3)
	9 (5)	13 (4)
		3 (5)

De waarde van de variabele, waarop de opgegeven frequenties betrekking hebben, is steeds tussen haakjes aangegeven.

De toetsing bij deze ongelijk verdeelde variabelen was als volgt:

Tabel 5. Toetsing van de verdelingsverschillen van de variabelen van tabel 4.

variabele nr.	toets	uitkomst
5	3x2 tabel	chi <sup>2</sup> = 6,76 0,02 < P < 0,05 *
6	2x2 tabel	chi <sup>2</sup> = 7,22 0,001 < P < 0,01 *

\* P kleiner dan de significantiegrens

12,13 en 14	3x2 tabel	$\chi^2=12,05$ $0,001 < P < 0,01^*$
13 en 14	2x2 tabel	$\chi^2=6,95$ $0,001 < P < 0,01^*$
38	Mann-Whitney U-toets	$z=-2,09$ $0,01 < P < 0,02^*$
67	Mann-Whitney U-toets	$z=-4,60$ $P < 0,00003^*$

De variabele 87 had zo'n grote variatie in waarden, dat het opgeven van frequenties niet zinvol is. De gemiddelde waarde van 87 in de PO groep bedraagt 179,8 en in de NO groep 315,5. De standaardafwijkingen bedragen resp. 47,2 en 46,7. Het verschil in gemiddelde tussen de twee groepen is groot genoeg om zonder meer de beslissing te nemen, dat de gemiddelde naonderzoekperiode van de PO groep belangrijk korter is dan die van de NO groep.

Betreffende de tertiusletsels (var. 38), dient nog te worden opgemerkt, dat er weliswaar relatief meer letsels van de malleolus tertius waren in de geopereerde groep (PO), doch dat de verdeling van de grootte van de aanwezige tertiusfragmenten bij beide groepen (PO en NO) niet duidelijk verschilde: ( $z=1,52$  en  $P=0,06$ ).

Om te beoordelen of er bij al deze verschillen nog wel enige vergelijking mogelijk is, werd nagegaan op welke wijze deze variabelen de toetsingskriteria beïnvloeden. Omdat de bespreking hiervan thans aan de orde is, moet ik vooruit lopen op de analyse van deze toetsingskriteria (zie hoofdstuk V C).

Ter beoordeling van de invloed van de variabelen werden kruistabellen gemaakt met de toetsingskriteria zowel in de totaal onderzochte groep, als in de PO en NO groepen. Als toetsingskriteria werden hiervoor gekozen de subjectieve scores 96 en 97, alsmede de objectieve somscore 175. Het resultaat was als volgt:

#### Variabelen 5 en 6.

Ten aanzien van variabele 5 was noch in de totale groep, noch in één der subgroepen PO en NO een verband tussen de variabele en één der toetsingskriteria aan te tonen. Ten aanzien van variabele 6 bleek in de NO groep een score  $\neq 0$  samen te hangen met een lagere (=gunstiger) score 175 ( $z=-3,14$ ). Dit zou doen verwachten, dat dezelfde samenhang bestaat tussen var. 6 en 96, dit is echter niet het geval ( $z=-1,02$ ). Als men op basis van deze verwachting éézijdig toetst blijkt een hogere score van 6 wel samen te hangen met een gunstiger score 97 (somscore subjectief welbevinden)  $z=-1,76$ .

#### Variabelen 12, 13 en 14.

Bij het beoordelen van de samenhang tussen deze variabelen, die het aantal A- resp. B- en C-fracturen in de PO en NO groepen weergeven, werd gezocht naar verschil in de score van de toetsingskriteria bij aanwezigheid van een A-, resp. B- of C-fractuur

en bij aanwezigheid van juist geen A-, resp. B- of C-fractuur. Deze verschillen werden steeds tweezijdig getoetst. Met uitzondering van het verschil tussen de scores van de toetsingscriteria bij C-fracturen en die bij juist geen C-fracturen. Volgens de opvattingen van Weber hebben de C-fracturen de slechtste prognose. Men zou dus verwachten, dat de aanwezigheid van een C-fractuur samengaat met een hogere score van de toetsingscriteria, dan de andere fractuurtypen. Daarom werd bij toetsing van dit verschil als  $H_0$ , de hypothese gebruikt, dat bij C-fracturen geen hogere (=ongunstiger) scores van de toetsingscriteria worden gevonden, dan bij de andere fracturen; op basis van deze hypothese werd éénzijdig getoetst.

Variabele 12 (de A-fracturen). Het bleek dat de aanwezigheid van een A-fractuur met een lagere objectieve somscore (175) samenging dan de aanwezigheid van een andere fractuur. Deze relatie was aantoonbaar zowel in de totale groep, als in de NO groep (resp.  $z=2.05$  en  $z=2.52$ ). In de PO groep kon een dergelijke relatie niet worden aangetoond. Bij geen van de groepen (Totaal, PO en NO), bleek de hoogte van de subjectieve scores samen te hangen met het wel of niet aanwezig zijn van een A-fractuur.

Variabele 13 (de B-fracturen). B-fracturen bleken in de NO groep samen te gaan met een hogere score 175 dan de andere fracturen ( $z=2.93$ ), dit geldt trouwens ook voor de C-fracturen ( $z=3.04$ ) en kan dus alleen veroorzaakt worden door het relatief gunstig beeld bij de resultaten van de A-fracturen ten aanzien van de objectieve somscore (175). In de totale groep kon de relatie tussen het aanwezig zijn van een B-fractuur en een ongunstiger score 175 echter niet worden aangetoond ( $z=1.84$ ). Ook binnen de PO groep kon deze relatie niet worden aangetoond.

Ten opzichte van de subjectieve scores (96 en 97) bleek er in geen van de drie groepen een aantoonbaar verschil te bestaan tussen de hoogte van deze scores bij de B-fracturen en die bij de andere fracturen.

Variabele 14 (de C-fracturen). Zoals boven reeds vermeld ging de aanwezigheid van een C-fractuur samen met een ongunstige objectieve score (175) in de NO groep. Dit was eveneens het geval binnen de totale groep ( $z=2.45$ ) en bij de PO groep ( $z=1.76$ ). Er zij aan herinnerd dat de verschillen bij de C-fracturen éénzijdig getoetst worden.

Ten opzichte van de subjectieve scores (96 en 97) blijkt de C-fractuur in alle groepen samen te gaan met een ongunstiger (hoger) score. Het is echter opvallend, dat juist in de NO groep de significantie minder sterk is (PO en NO bij 97 is resp.  $z=2.24$  en  $z=1.75$ ).

In het voorgaande werden steeds fracturen van een bepaald type vergeleken met alle andere fracturen samen. Bijvoorbeeld A-fracturen met 'niet A'-fracturen.

In symbolen kan men dat uitdrukken als A met  $\bar{A}$ , B met  $\bar{B}$  enz. Voor de overzichtelijkheid werden de besproken vergelijkingen in het volgende schema ondergebracht. Hierin worden de waarden van de medianen in de diverse groepen aangegeven. Indien bij een vergelijking een significant verdelingsverschil werd ge-

vonden wordt dit met een → aangegeven. De richting van de pijl geeft de groep met de gunstigste verdeling aan.

	A				B				C			
	PO		NO		PO		NO		PO		NO	
	A	A	A	A	B	B	B	B	C	C	C	C
96	0	0	0	0	0	0	0	0	0 → 0		0 → 0	
97	0	0	0	0	0	0	0	0	0 → 0		1 → 0	
175	4,5	7	4 ←	7,5	6	7	9 →	5	8 → 6		11 → 5	

In de subjectieve scores blijken de meeste medianen de waarde nul te hebben. Dit was te verwachten aangezien in iedere groep meer dan de helft van de patiënten subjectief geheel goed wordt, zodat de mediaan dan op nul komt te liggen.

Variabele 38 (Tertius fragment) bleek in alle groepen samen te hangen met het resultaat en wel zodanig, dat de aanwezigheid van een tertiusfragment een ongunstiger score van alle toetsingscriteria tengevolge heeft.

Variabele 67 (Primaire dislocatie). Om de samenhang van deze variabele met de toetsingscriteria na te gaan werden rangcorrelatiecoëfficiënten berekend. Hierbij geeft een hogere coëfficiënt een nauwere samenhang weer dan een lagere. Deze coëfficiënten worden in onderstaande tabel weergegeven in duizendsten.

Tabel 6. Rangcorrelatiecoëfficiënten (Spearman) in duizendsten tussen de variabele 67 enerzijds en de variabelen 96, 97 en 175 anderzijds. Ze zijn berekend in de drie steekproeven: totale groep onderzochte patiënten, de PO groep en de NO groep.

Var.nr.	Totaal	PO	NO
67-96	101	228*	144
67-97	201*	424*	205*
67-175	485*	356*	654*

Uit de tabel blijkt, dat variabele 67 binnen de totale en de NO groep vooral met de objectieve score (175) samenhangt. De samenhang met variabele 96 (de tevredenheid van de patiënt) is in alle groepen minder duidelijk. Binnen de PO groep valt de sterke samenhang met de subjectieve somscore (97) op.

In de NO groep en de totale groep kon geen verband tussen 67 en 96 worden aangetoond.

Variabele 87, het tijdsverloop tussen het optreden van de fractuur en het verzamelen van de naonderzoekgegevens wordt afzonderlijk in het volgende hoofdstuk besproken. Op deze plaats is relevant dat geen sterke samenhang van deze variabele met de toetsingscriteria 96, 97 en 175 kon worden aangetoond.

\* P Kleiner dan de significantiegrens.

Wat moeten we uit dit alles besluiten omtrent de vergelijkbaarheid van de twee groepen PO en NO?

Bij alle besproken variabelen met uitzondering van 5, 6 en 12 werd een ongunstige invloed vastgesteld op de uiteindelijke beoordelingscriteria. Met dien verstande dat een hogere waarde van de betrokken variabele het optreden van ongunstiger waarden van de beoordelingscriteria in de hand werkt. De uitzonderingen hierop:

5 (zwaarte van de werkkring) leek geen invloed uit te oefenen op de waarde van de toetsingscriteria, terwijl bij het aanwezig zijn van relatief veel patiënten met zwaardere vrije tijdsbelasting (6) of van patiënten met een A-fractuur (12) de beoordelingscriteria juist in gunstige zin worden beïnvloed.

In de tweede plaats viel het bij de bespreking van de samenhang tussen variabele 67 en de toetsingscriteria op dat de subjectieve scores zich anders leken te gedragen, dan de objectieve somscore. Dit fenomeen zal nog meerdere malen signaleerd worden.

Indien we nu terugkeren tot de verdelingsverschillen van deze variabelen tussen de PO en NO groepen, dan blijkt dat de verschillen steeds ten ongunste van de PO groep uitvallen, d.w.z. de PO groep bevat relatief minder van de gunstige A-fracturen, relatief meer van de ongunstige tertiusfragmenten, meer ongunstige dislocaties en als enige gunstige uitzondering bevat de PO groep meer van de gunstige zware vrije tijdsbelastingen. Van dit laatste mag echter geen oorzakelijke relatie met een gunstig behandelingsresultaat verwacht worden, hoogstens zal een gunstig resultaat de patiënt in staat stellen er een meer belastende vrije tijdsbesteding op na te houden.

Dit wil zeggen, dat de vergelijking van de twee groepen PO en NO slechts in één richting mogelijk is:

Zou bij de bewerking blijken, dat bij het naonderzoek de PO groep slechter uitvalt, dan betekent dit niet dat de operatie een slechter resultaat tengevolge heeft gehad, doch dat zij niet in staat is geweest om het relatief hoog aantal ongunstige factoren, dat aanwezig is, bij de patiënten in deze groep, te nivelleren. Zou daarentegen de PO groep er alsnog beter afkomen dan de NO groep, dan lijkt de conclusie gerechtvaardigd dat het behandelingsverschil hier een bepalende rol heeft gespeeld.

#### **B *Zijn de groepen representatief?***

Hoewel het interessant kan zijn om conclusies over de onderzochte groepen te trekken, is het de bedoeling het inzicht betreffende de voorkeursbehandeling bij patiënten met een enkelfractuur te verruimen. Hiertoe dient de generalisatie vanuit de steekproef naar de algemene groep patiënten gerechtvaardigd te zijn. Strikt genomen is deze generalisatie niet meer mogelijk indien het aselechte karakter van de steekproef ten opzichte van de algemene groep heeft geleden. In hoofdstuk III werd reeds betoogd dat, zelfs bij een aanvankelijk complete steekproef, tijdens het naonderzoek een verschuiving is opgetreden ten opzichte van de patiënten waarvoor

de conclusies relevant worden geacht. Door overlijden is de naonderzochte groep gemiddeld jonger geworden dan de algemene patiënt met een enkelfractuur, terwijl ook door verhuizing een aantal patiënten niet voor naonderzoek bereikbaar is. Omdat werd verondersteld, dat deze verschuivingen niet door de enkelfractuur werden veroorzaakt, werd de leeftijd gebruikt om de invloed op de steekproef te beoordelen:

Tabel 7. Gemiddelde leeftijden.

	PO groep Gem. leeftijd	NO groep Gem.leeftijd	P (t-toets)
Onderzochte groep	42,6 (85)	39,2 (156)	0,1 < P < 0,2
Overleden	59,0 (5)	62,4 (21)	—
Vertrokken	41,7 (7)	35,7 (27)	—

In deze tabel zijn de frequenties tussen haakjes vermeld.

De in de tabel vermelde gemiddelde leeftijden hebben betrekking op de leeftijd ten tijde van het ongeval. Het zijn dus inderdaad de oudere patiënten, die door overlijden niet konden worden onderzocht. De leeftijd van de vertrokken patiënten wijkt niet duidelijk af van die van de onderzochte patiënten. Indien de leeftijd een relatie heeft met de beoordelingscriteria, dan geeft de onderzochte steekproef een onjuist beeld over het te verwachten resultaat bij een willekeurige patiënt met een enkelfractuur omdat een onevenredig groot aantal oudere patiënten niet is onderzocht.

Om dit te onderzoeken werd naar samenhang gezocht tussen de leeftijd van de patiënt (var. 4) en de toetsingscriteria (96, 97 en 175), hiertoe werden rangcorrelatiecoëfficiënten berekend:

Tabel 8. Rangcorrelatiecoëfficiënten (Spearman) tussen de variabele 4 enerzijds en de variabelen 96, 97 en 175 anderzijds. De coëfficiënten worden gegeven in duizendsten.

Var.nr.	Totaal	NO	PO
4 met 96	33	37	81
4 met 97	161*	90	399*
4 met 175	295*	373*	206*

Indien de uit deze coëfficiënten berekende t-verdeelde grootheid tot het vinden van een overschrijdingskans kleiner dan de significantiegrens leidde, werd dat in de tabel met een \* aangegeven.

Het valt op dat de leeftijd samen blijkt te gaan met de objectieve somscore (175) in alle drie de groepen, doch, dat er een veel onduidelijker samenhang bestaat tussen de leeftijd van de patiënten en de subjectieve scores. Wederom dus een ander gedrag van deze twee toetsingscriteria. Tenslotte blijkt een hogere leeftijd in de meeste

\* P Kleiner dan de significantiegrens.



gevallen samen te gaan met een hogere score van de toetsingscriteria zodat door de selectie, die optreedt doordat een aantal oudere patiënten niet voor naonderzoek kan worden bereikt, bij de naonderzochte groep een behandelingsresultaat wordt gevonden, dat gunstiger is, dan hetgeen, dat bij een willekeurige groep patiënten, die behandeld dienen te worden voor een enkelfractuur, verwacht mag worden.

Zoals reeds werd vermeld op pag. 78 onttrokken zich 41 patiënten aan het naonderzoek. Meestal door niet reageren op het verzoek te willen komen voor een naonderzoek ondanks minstens 3x een schriftelijk verzoek nadat het adres bij de burgerlijke stand was geverifieerd. Al deze patiënten behoorden tot de NO groep.

Bij deze patiënten was het denkbaar, dat hun grote ontevredenheid met het resultaat van de behandeling van hun enkelfractuur of het feit dat ze inmiddels elders voor complicaties waren behandeld tot hun onwillige houding aanleiding gaf. Met andere woorden, dat ze een speciaal slechte groep vertegenwoordigden. Om deze reden werd een steekproef van 10 patiënten getrokken, die door middel van huisbezoek en persoonlijk vervoer toch naar het ziekenhuis werden gebracht teneinde adequate röntgencontrole mogelijk te doen zijn. Het weergeven van alle 175 gegevens van deze tien patiënten zou wel veel ruimte innemen en de overzichtelijkheid niet ten goede komen. Het type fractuur, de gemiddelde leeftijd, het geslacht, de primaire dislocatie, maar vooral de bevindingen bij onderzoek werden als criterium gebruikt om te beoordelen, of de uitval van deze patiënten de conclusies uit het naonderzoek sterk beïnvloedde. De gegevens zijn samengevat in tabel 9.

Tabel 9. Frequenties van variabelen ter vergelijking van de onderzochte NO groep en een steekproef van tien patiënten. Deze steekproef werd getrokken uit een groep van 41 patiënten, die weigerden voor naonderzoek te komen.

	Onderzochte NO groep	Steekproef van 10 patiënten
Mannen	93	8
Vrouwen	63	2
Gem.leeftijd	39,2	42,2
Aantal A-fract.	23	1
Aantal B-fract.	74	8
Aantal C-fract.	21	1
Primaire dislocatie	32x0; 31x1; 45x2 24x3; 13x4; 3x5	4x0; 3x1; 3x2
Patiënt tevreden (var.96)	124x0; 28x1; 4x2	9x0; 1x1
Subject.somscore (var.97)	82x0; 32x1; 12x2 13x3; 6x4; 5x5; 2x6 2x8 en 2x9	8x0; 1x1; 1x2
Objectieve somscore (var.175)	gem. 9,82	gem. 10,31

Op grond van deze gegevens bleek er geen aanleiding te bestaan voor de vrees dat de

niet verschenen patiënten een speciale groep vertegenwoordigden wat betreft hun enkelfractuur, of hun klachten daarvan.

Wel valt het geringe aantal vrouwelijke patiënten op, doch het was soms moeilijk om de patiënten er toe te bewegen tijdens de normale werkuren voor het onderzoek mee te gaan. Beide vrouwen uit deze groep hadden een normale werkkring en waren alleen bereid om na werktijd naar het ziekenhuis te komen. Deze reden om niet voor naonderzoek te komen zal relatief voor meer mannen dan vrouwen gelden. Angst voor het ziekenhuis speelt ook een rol. Volgens een telefonisch contact, zou één der wat oudere patiënten uit de steekproef sinds de enkelfractuur zo slecht ter been zijn, dat hij nauwelijks de kamer uit kwam. Dit volgens zijn dochter bij wie hij inwoonde. Toen ik onaangekondigd bij patiënt kwam, bleek hij niet thuis te zijn, maar was met zijn kleinzoon van 12 jaar aan het voetballen op een naburig speelveldje! Bij het onderzoek bleek hij van de enkelfractuur geen last te hebben en ook de röntgenfoto was vlekkeloos, doch hij bekende wel dat hij zo'n hekel had aan het ziekenhuis dat hij nooit uit zichzelf zou zijn gekomen.

Ten aanzien van de representativiteit van de onderzochte groepen kunnen we dus het volgende vaststellen:

Onvermijdelijk verlies ontstaat door tussentijds overlijden van patiënten, het heeft niet te maken met de enkelfractuur, doch doet een verschuiving optreden ten opzichte van de gemiddelde leeftijd van de onderzochte groep, die daarom niet meer geheel representatief is voor een willekeurige patiënt.

Onvermijdelijk verlies ontstaat eveneens door het buiten bereik verhuizen van patiënten. Aangezien het maken van goede standaard röntgenfoto's en eventueel speciale foto's voor het vaststellen van bandinsufficiënties als voorwaarde werd gesteld voor een adequaat onderzoek, werd een patiënt als onbereikbaar gezien indien hij naar een andere provincie of ander land was verhuisd.

Veronderstellend dat de verhuizing niet met de enkelfractuur te maken heeft en nadat vastgesteld werd dat ook de leeftijd niet duidelijk anders was dan van de onderzochte groep, wordt dit verlies als een toevallig verlies beschouwd dat de representativiteit niet schaadt.

Tenslotte ontstaat verlies doordat een aantal patiënten niet wil komen. Het aselechte karakter van dit verlies staat niet zonder meer vast. Doch uit een steekproef bleek dat ook bij deze groep geen duidelijk verschil is te vinden met de ernst van de klachten van de onderzochte groep, terwijl ook de leeftijdsverdeling niet afweek van die bij de onderzochte patiënten. Ook hier kon dus gesteld worden, dat het weg-blijven waarschijnlijk met factoren te maken heeft, die niet of nauwelijks met de enkelfractuur samenhangen en kan het verlies als toevallig worden beschouwd.

### C Toetsingskriteria

Om het resultaat te beoordelen en het gemiddeld resultaat van de groepen te kunnen vergelijken moet een toetsingskriterium worden gebruikt. De criteria, die door Weber worden gehanteerd werden reeds beschreven alsmede de redenen om naar een ander kriterium te zoeken.

Op pag. 51 werd betoogd, dat het resultaat beoordeeld moet worden naar de subjectieve mededelingen van de patiënt, terwijl de objectieve bevindingen slechts dienen om te kunnen beoordelen hoe de patiënt zich in de toekomst zal voelen, met andere woorden de objectieve bevindingen dienen er toe om een prognose te kunnen formuleren over het subjectieve beleven van de patiënt. Naar aanleiding van een dergelijke prognose kan worden beoordeeld of b.v. een therapeutische handeling de prognose kan verbeteren.

Aangezien de exacte relatie van de objectieve bevindingen en de prognose niet bekend is, is het onmogelijk om de objectieve bevindingen in een subjectieve relatie om te zetten waarmee een relevante beoordeling in één cijfer mogelijk zou worden. Daarom werden aan de beoordelingskriteria van Weber drie andere toegevoegd:

- subjectieve tevredenheid (variabele 96)
- subjectieve somscore, als totaal van de variabelen 88, 89, 90 en 96 (variabele 97)
- objectieve somscore, als totaal van variabelen 98 t/m 102, 104, 109, 111, 113 en 115 t/m 173 (variabele 175)

Zo hebben we te maken met 5 toetsingskriteria:

- 1) Weber somscore (var. 94)
- 2) Weber eindscore (var. 95)
- 3) Subjectieve tevredenheid (var. 96)
- 4) Subjectieve somscore (var. 97)
- 5) Objectieve somscore (var. 175)

Hiervan heeft var. 96 een duidelijke kwalitatieve interpretatie, zijn rangorde werd reeds besproken in hoofdstuk IV.

De somscores hebben echter een onduidelijke kwalitatieve interpretatie.

Ze zijn ontstaan door het bij elkaar tellen van een aantal variabelen, die min of meer bij elkaar horen, maar wat hun waarde te betekenen heeft is niet precies vast te stellen.

Variabele 95 stelt een samentrekking van variabele 94 voor, zodat hierbij helemaal niet meer duidelijk wordt wat de kwalitatieve inhoud van deze variabele is.

Somscores, moeten aan een aantal voorwaarden voldoen om als een redelijk toetsingskriterium te kunnen gelden:

- De samenstellende onderdelen moeten onderling een redelijke relatie hebben, die niet al teveel uit elkaar loopt
- De samenstellende delen, moeten allen ongeveer dezelfde relatie bezitten met de somscore.

Om dit te kunnen beoordelen werden correlatietabellen samengesteld van de somscores en hun samenstellende onderdelen:

### 1. Weber somscore/eindscore

Tabel 10\*. Correlatiecoëfficiënten (in duizendsten) berekend over de totaal onderzochte groep patiënten.

Pijn	(88)	1000											
Lopen	(89)	301	1000										
Activ.	(90)	473	444	1000									
Anatomie	(91)	198	285	268	1000								
B.S.G.	(92)	333	360	342	315	1000							
O.S.G.	(93)	264	393	356	229	490	1000						
Somscore	(94)	567	645	661	709	692	655	1000					
Eindscore	(95)	396	336	284	760	451	352	722	1000				
Subj.tevr.	(96)	580	433	723	277	386	360	639	295	1000			
Subj.somsc.	(97)	733	706	812	318	433	422	779	416	841	1000		
↑ Variabelen nr. →		88	89	90	91	92	93	94	95	96	97		

Het valt hierbij op, dat de drie, voor een belangrijk deel subjectief bepaalde, variabelen pijn, lopen en activiteit onderling sterker samenhangen (hogere correlatiecoëfficiënten) dan met het objectieve gegeven: de anatomische verhoudingen op de röntgenfoto (var. 91). Als dit even terzijde wordt gesteld, dan blijkt de somscore van Weber zich redelijk te verhouden ten opzichte van zijn samenstellende variabelen, hoewel hier een wat nauwere samenhang met de anatomie lijkt te bestaan dan met de meer subjectieve gegevens. Ter vergelijking werd de eindscore van Weber eveneens in de tabel opgenomen. Deze blijkt veel te sterk met variabele 91 samen te hangen, deze correlatiecoëfficiënt is ongeveer even hoog als die met de somscore waar hij van is afgeleid. Als toetsingskriterium is hij daarmee onbruikbaar, als het er om gaat ook de betekenis van de variabelen 88, 89 en 90 in dit criterium weer te geven.

Indien één eindscore wordt vastgesteld, dan moet daarvan worden geëist, dat deze de subjectieve bevindingen van de patiënt goed weergeeft. Het strikt objectieve gegeven 91 blijkt echter wel sterk met de Weber eindscore en de Weber somscore samen te hangen, doch opvallend laag te correleren met de subjectieve tevredenheid

\* In feite zijn de interpretaties verricht aan de hand van correlatiecoëfficiënten, die werden berekend binnen de PO en NO groepen. Deze gaven echter geen wezenlijk ander beeld dan de coëfficiënten die over de hele groep berekend werden. Zodat ter wille van de eenvoud wordt volstaan met de presentatie van deze tabel.

(96). De grote invloed van de anatomie (91) op de Weber eindscore (95), kan mogelijk verantwoordelijk gesteld worden voor het relatief laag correleren tussen 95 en de subjectieve score 96. Het lijkt alsof de subjectieve gegevens zich anders gedragen dan de objectieve, onderling correleren zowel de subjectieve (88, 89, 90, 96) en de objectieve (91, 92, 93) ongeveer even sterk, maar tussen subjectieve en objectieve scores treden steeds de laagste correlaties op. Dit komt het duidelijkst tot uiting bij de subjectieve somscore. De laagste correlaties treden ten opzichte van de objectieve variabelen op, terwijl de correlaties ten opzichte van de andere gegevens steeds bijna twee keer zo hoog zijn. Dit gevoegd bij het feit dat de relatie tussen de objectieve score en de prognose van het subjectief welbevinden onbekend is, geeft aanleiding tot de volgende opvatting:

- 1) Het weergeven van het naonderzoekresultaat in één enkele score is onjuist.
- 2) Hoewel de samenstellende variabelen van de Webersomscore redelijk met elkaar en met de somscore correleren lijkt de Webersomscore daarom toch geen juiste weergave te kunnen geven van het resultaat van de behandeling van een enkel-fractuur.
- 3) De objectieve afwijkingen bij patiënten die destijds wegens een enkelfractuur werden behandeld hangen sterker met elkaar dan met het subjectieve welbevinden van de patiënten samen.
- 4) De subjectieve klachten bij patiënten die destijds werden behandeld wegens een enkelfractuur, hangen sterker met elkaar dan met de objectieve afwijkingen samen.
- 5) Om een hanteerbare eindbeoordeling te verkrijgen van naonderzoekpatiënten verdient het de voorkeur om twee eindscores te vormen:
  - a) een somscore, die de subjectieve gegevens samenvat en
  - b) een somscore, die de objectieve bevindingen samenvat.

In de bewerkingen van het onderzoekmateriaal werden hiervoor gekozen:

- a) De subjectieve somscore, verkregen uit de variabelen 88, 89, 90 en 96 (=variabele 97).
- b) De objectieve somscore als totaal van de variabelen 98 t/m 102, 104, 109, 111, 113 en 115 t/m 173 (=var. 175).

Om na te gaan hoe de zuiver subjectieve tevredenheid van de patiënt zich in verhouding tot een aantal variabelen gedraagt zal naast deze toetsingscriteria soms ook variabele 96 worden gebruikt.

## 11. *De subjectieve somscore (var. 97)*

Hoewel de correlaties van de samenstellende variabelen van deze score reeds in tabel 10 zijn vermeld worden ze ter wille van de duidelijkheid hier nog eens apart opgegeven:

Tabel 11. Correlatiecoëfficiënten van de samenstellende variabelen van de subjectieve somscore 97 (in duizendsten).

Pijn	(88)	1000				
Lopen	(89)	301	1000			
Activiteit	(90)	473	444	1000		
Subj.tevredenheid	(96)	580	433	723	1000	
Subj.somscore	(97)	733	706	812	841	1000
Variabelen		88	89	90	96	97

### III. De objectieve somscore (var. 175)

Het opnemen van een correlatietabel, die uit 68 variabelen is samengesteld, lijkt de duidelijkheid niet ten goede te komen.

Hij werd echter wel berekend, en het bleek dat de röntgegegevens in het algemeen onderling hoog correleerden evenals de gegevens, die afkomstig zijn van het beoordelen van de weke delen verhoudingen. Deze twee groepen bleken echter relatief duidelijk lager met elkaar te correleren.

Om na te gaan of bepaalde variabelen zich min of meer groepsgewijs opstellen werd een factoranalyse\* uitgevoerd. Na rotatie van 5 van de 10 ingevoerde factoren (varimax criterium), traden factorladingen op, die bij 4 van de 5 factoren wel enige betekenis hadden.

#### Factor 1:

Deze factor wordt uitsluitend gevormd door röntgen gegevens.

	factorladingen .
Spleet in het gewrichtsvlak (130)	551
Rest luxatiestand naar lateraal (131)	538
Sclerose mediaal en lateraal (138 & 139)	673&676
Gewrichtsspleetvernauwingen (140 t/m 144 en 146)	690, 715, 700, 707, 556 en 663
Gewrichtsvlakonregelmaticgheden (147 t/m 150 en 153)	656, 713, 734, 576, 564
Osteophyt vorming aan de voorzijde (154)	598
Verstoorde botstructuur van de tibia (158)	596

Het valt hierbij op dat bij de vernauwingen de relatie tussen mediale malleolus en de talus (145) blijkbaar een minder grote rol speelt, terwijl bij de onregelmaticgheden de relatie tussen talus en mediale malleolus zowel als laterale malleolus ontbreekt (151 & 152).

#### Factor 2:

Deze factor wordt gevormd door weke delen gegevens en röntgegegevens, die niet met het röntgenaspect van de gewrichtsspleet zelf te maken hebben:

\* In feite werden de interpretaties verricht aan de hand van factoranalyses binnen de PO en NO groepen. Het presenteren van al deze gegevens zou de duidelijkheid niet ten goede komen. Aangezien geen wezenlijke verschillen bestonden worden alleen gegevens betreffende de totaal onderzochte patiëntengroep gepresenteerd.

	factorladingen
Pigmentaties (111)	485
Klachten bij slecht weer (117)	422
Weefsel praeachillair (119)	593
Valgusstand (136)	-442
Verkalking aan de achterzijde (167)	404

### Factor 3:

Hierin spelen röntgegegevens geen enkele rol:

	factorladingen
Cysteuze zwelling over de voorste syndesmose (115)	503
Drukpijn op laterale band (121)	430
Drukpijn op de voorste syndesmose (122)	630
Drukpijn op de achterste syndesmose (123)	461

### Factor 4:

Drie röntgegegevens, die met een zeer lage frequentie voorkomen:

	factorladingen
Restdislocatie mediale tibiahoek (128). Dit komt in het totale materiaal 17 maal voor	652
Krachtlijnen malleolus lateralis (161), hetgeen 2 maal voorkomt	559
Krachtlijnen malleolus medialis (163), hetgeen slechts 1 maal voorkomt.	644

Door de in het algemeen geringe genuanceerdheid van de gegevens en de lage frequenties van de afwijkingen zijn deze bevindingen moeilijk op hun betekenis te schatten en vooral aan de laatste factor mag weinig waarde worden gehecht. Toch wordt de indruk ondersteund, dat de röntgegegevens zich tamelijk geïsoleerd van de weke delen afwijkingen opstellen zodat het de vraag wordt of deze gegevens wel binnen één somscore thuis horen. Het leek aantrekkelijk om van deze groepen drie aparte objectieve somscores te maken en op hun gedrag te analyseren. Het gevolg zou echter zijn dat er tenslotte bij iedere patiënt vier eindcijfers zouden ontstaan, die het resultaat weergaven. Dit zou de onderlinge vergelijking van de groepen vrijwel onmogelijk maken, zodat hier tenslotte van werd afgezien en toch gekozen werd voor het hanteren van een totale somscore van alle objectieve gegevens. Bij de bewerking zal echter blijken, dat de weke delen zich niet alleen anders gedragen binnen de somscore, doch dat ook hun relatie met het subjectief welbevinden sterker verschilt van die van de röntgegegevens dan wenselijk is voor variabelen, die binnen één eindscore worden verzameld.

## Hoofdstuk VI

### DE FACTOR TIJD: DE BETEKENIS VAN HET TIJDSVERLOOP TUSSEN FRACTUUR EN NAONDERZOEK

Bij de bespreking van de vergelijkbaarheid van de groepen PO (primair geopereerde patiënten en NO (niet geopereerde patiënten), is de analyse van het grote verschil in gemiddeld tijdsverloop tussen fractuur en naonderzoek achterwege gebleven.

Het is zo groot, dat in principe tot onvergelijkbaarheid zou moeten worden besloten. Alleen indien aantoonbaar is dat het tijdsverschil geen of slechts een zeer geringe invloed heeft op de beoordeelde variabelen, kan een vergelijking van de twee groepen tot conclusies leiden, die los staan van het tijdsverschil. Om na te gaan hoe de invloed van de tijd op de verschillende variabelen was, werd getracht door middel van grafische voorstellingen, waarbij de variabelen tegen de tijd werden uitgezet, een indruk te krijgen over een mogelijke samenhang. Vervolgens werd een nieuwe variabele gedefinieerd (174) die de followup periode (var. 87) in klassen van 52 weken weergaf en wel als volgt:

Variabele:	174	87
	0	0–130
	1	131–182
	2	183–234
	3	235–286
	4	287–338
	5	339–390
	6	≥391

Nu werden de frequenties bepaald van variabele 174 binnen de groepen PO en NO:

Tabel 12. Verklaring in de tekst.

	PO	NO
174=0	17	—
174=1	19	1
174=2	40	—
174=3	9	39
174=4	—	74
174=5	—	34
174=6	—	7



Hierbij bleek er een klasse te bestaan waarbij de followup periode bij de PO en NO groep gelijk was binnen de grenzen van de klasse. Deze groep werd afzonderlijk beoordeeld, in de hoop dat er ondanks de kleine aantallen vooral van de PO groep conclusies uit te trekken waren. Deze groep is in de tabel omlijst.

Aangezien er zo'n geringe overlapping is tussen de groepen PO en NO wat betreft de tijd heeft het weinig zin om relaties tussen de tijd en diverse onderzochte variabelen van de totaal onderzochte patiëntengroep na te gaan. Immers ieder verschil dat tussen de PO en NO groep bestaat zal zich bij bewerken van de totale groep ten aanzien van de tijd ook gedragen als afhankelijk van deze tijd.

De invloed van de tijd op de variabelen moet dan ook gezocht worden binnen de twee groepen zelf, hiertoe werden correlatiecoëfficiënten bepaald tussen de tijd (var. 87) en een aantal subjectieve en objectieve variabelen.

### **A De grafische voorstellingen**

Omdat grote hoeveelheden cijfers minder direct aanspreken dan grafische voorstellingen, werden, als eerste oriëntatie, de onderzoekvariabelen tegen de tijd (var. 87) uitgezet.

De mogelijkheden om duidelijke grafische voorstellingen te krijgen werden bemoeilijkt door:

- 1) Het relatief kleine aantal afwijkingen, waardoor de meeste variabelen de waarde nul hebben.
- 2) Het geringe discriminerend vermogen bij het bepalen van de waarde van de meeste variabelen waardoor geen tussenwaarden ontstaan doch steeds natuurlijke getallen.
- 3) De geringe schaalbreedte van de meeste variabelen, waardoor de variabelen slechts een gering aantal waarden kunnen hebben.

Hierdoor vallen vele punten van de grafieken samen, ontstaan vele niveauwisselingen als getracht wordt een vloeiende grafiek te doen ontstaan, zijn de mogelijkheden tot 'smoothing' beperkt en is ruisonderdrukking vrijwel onmogelijk. Toch werd met behulp van de computer getracht grafische voorstellingen (= plots) te maken, die met een cijfer aangeven hoeveel punten op een bepaalde plaats samenvallen. Indien meer dan 9 punten samenvallen wordt dit met een + aangegeven.

De figuren 6 t/m 11 zijn voorbeelden van deze plots.

Ongeveer midden over de plots is met twee verticale lijnen het in de tijd overlappende gebied van de PO en NO groepen aangegeven. Rechts van deze lijnen hebben de aangegeven punten dus uitsluitend betrekking op patiënten uit de NO groep, terwijl links van deze lijnen nog slechts één NO patiënt thuis hoort (bij 177 weken), voor de rest hebben alle punten uitsluitend op patiënten uit de PO groep betrekking.

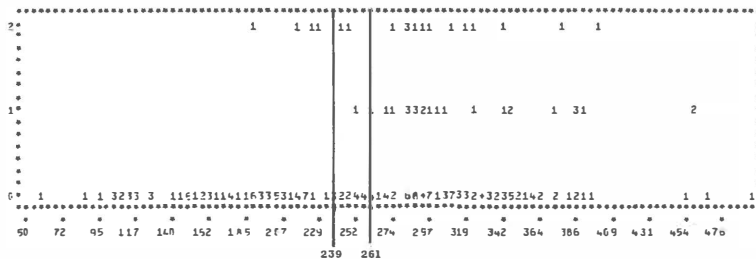


Fig. 6 Plot van var. 87 en var. 115 (Cysteuse zwelling over voorste syndesmose)

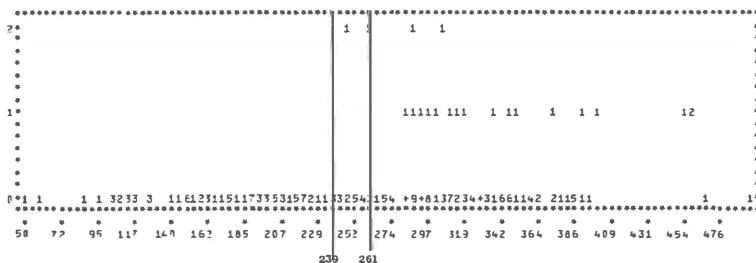


Fig. 7 Plot van var. 87 en var. 126 (Restdislocatie van de laterale malleolus)

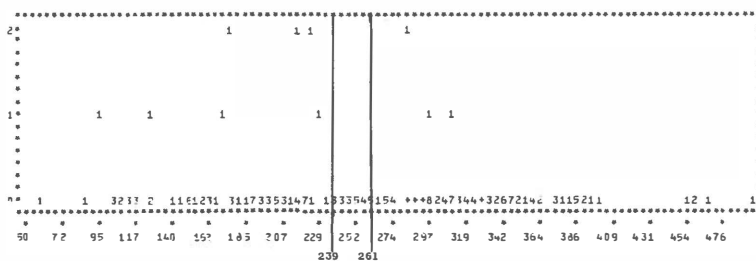
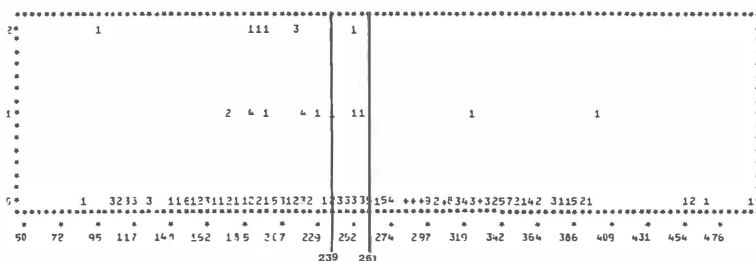


Fig. 8 Plot van var. 87 en var. 109 (Pathologische vaattekening)



**Fig. 9** Plot van var. 87 en var. 119 (Weefsel praeachillair)

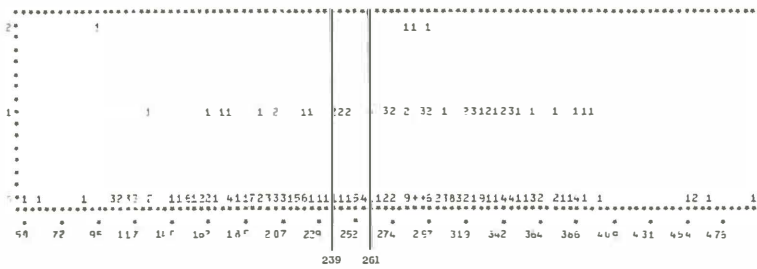


Fig. 10 Plot van var. 87 en var. 144 (Vernauwing van de gewrichtsspleet tussen malleolus lateralis en talus)

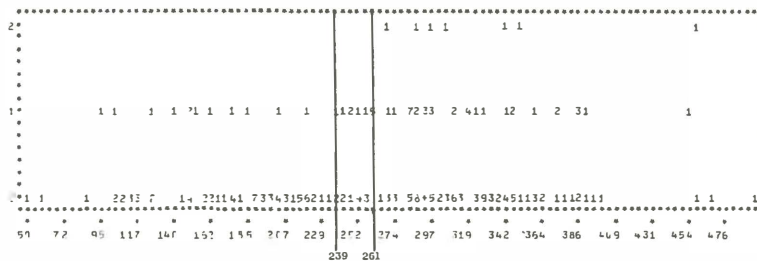


Fig. 11 Plot van var. 87 en var. 153 (Onregelmatig gewrichtsvlak tussen tibia en fibula)

Bekijken van de plots kan tot de volgende conclusies leiden:

- 7) Verreweg de meeste variabelen zitten op de nullijn.
- 2) De voorstellingen kunnen suggestief een verschil tussen de PO en NO groep weergeven. Zo lijkt er nogal wat verschil te bestaan tussen links en rechts van de vertikale lijnen bij de plots van de variabelen 126 (restdislocatie laterale malleolus) 115 (Cysteuse swelling over de voorste syndesmose), beiden ten nadele van de NO groep en 109 (homolaterale pathologische vaattekening) ten nadele van de geopeerde groep (PO). Binnen de benadeelde groep, dus rechts, respectievelijk links van de verticale lijnen, is er echter geen duidelijke invloed van een kortere of langere naonderzoekperiode te zien.
- 3) Een aantal figuren geven noch een tijdinvoel links of rechts van de vertikale lijnen te zien, noch een verschil tussen links en rechts: b.v. de plots van de variabelen 144 en 153.
- 4) Een enkele lijkt zowel een duidelijk verschil tussen links en rechts te demonstrenen (var. 119) als een invloed van de tijd binnen de benadeelde groep, in dit geval links, dus de PO groep.

De toetsings bevindingen op de exacte frequenties van de gesignaleerde links/rechts verschillen, dus verschillen tussen de PO en NO groep is als volgt:

Variabele nummer	Toets	Uitkomst
109	2x2 tabel	$\chi^2 = 4,01$ $0,02 < P < 0,05$
115	2x3 tabel	$\chi^2 = 19,16$ $P < 0,001$
119	2x2 tabel	$\chi^2 = 37,03$ $P < 0,001$
126	2x2 tabel	$\chi^2 = 12,20$ $P < 0,001$
144	Fisher (exact)	$P = 0,03$
153	2x2 tabel	$\chi^2 = 6,96$ $0,001 < P < 0,01$

Hoewel dus bij de toetsing bleek dat er ook significante verschillen tussen de groepen PO en NO bestonden wat betreft de verdeling van de variabelen 144 en 153, terwijl dit bij de plots niet duidelijk naar voren was gekomen, zijn de overschrijdingskansen flink wat groter dan bij de meeste andere variabelen. Hier blijkt uit dat de plots weliswaar minder onderscheidend vermogen verschaffen dan een toetsing, zoals natuurlijk te verwachten is, maar dat zij toch wel degelijk valide informatie kunnen geven.

Een overtuigende invloed van de tijd (var. 87) werd echter op geen enkele plot waargenomen en hoewel dit geen bewijs is dat de tijdsinvloed nu in het onderzoek verwaarloosbaar is, kunnen we dit wel als een voorlopige geruststelling beschouwen.

#### B De overlappende tijdsklasse (var. 174= 3)

Zoals reeds werd besproken is het mogelijk twee groepen patiënten uit de PO-respectievelijk uit de NO groep te vergelijken, waarvan het tijdsverloop tussen fractuur en naonderzoek steeds in dezelfde tijdsklasse valt (var. 174=3). Indien bij deze twee groepen significante verdelingsverschillen kunnen worden aangetoond bij bepaalde variabelen, dan is dat een ondersteunend argument voor de bewering dat indien dezelfde verdelingsverschillen worden gevonden bij vergelijking van de totale PO en NO groepen deze verschillen niet (alleen) door het tijdsverschil zijn veroorzaakt.

Vergeleken en getoetst werden de variabelen 96 t/m 167, 171 en 175. Het is weinig zinvol om alle frequenties en toetsingen bij deze twee groepen te vermelden. Vaak waren de frequenties zo klein dat een zinvolle toetsing niet mogelijk was, doch er bleken toch een aantal duidelijke verschillen te bestaan:

De PO groep bleek significant BETER uit te vallen in de verdeling van de variabelen:

Variabele nummer	Toets	Uitkomst
<u>97</u>	Mann Whitney	$z = 1,95$ $P = 0,026$
115	Mann Whitney	$z = 2,08$ $P = 0,019$
166	Mann Whitney	$z = 1,93$ $P = 0,027$

De geopereerde groep (PO) bleek echter SLECHTER uit te vallen in de verdelingen van de variabelen:

Variabele nummer	Toets	Uitkomst
111	Mann Whitney	$z = 3,68$ $P < 0,00016$
118	Mann Whitney	$z = 2,97$ $P = 0,0014$
119	Mann Whitney	$z = 4,30$ $P < 0,00003$

Het subjectieve toetsingskriterium (97) viel dus significant gunstiger voor de PO groep uit, het is in de tabel onderstreept.

Het is van belang op te merken, dat de variabelen, die bij de PO groep ongunstiger waren verdeeld, twee waarnemingen aan de weke delen betreffen (111 en 119) en het evenmin aan de röntgenbevindingen gebonden startpijnfenomeen (118). Bevindingen in de zelfde richting bij vergelijking van de totale PO en NO groepen kunnen dus kennelijk niet door het tijdsverschil tussen de twee totale groepen worden verklaard en mogen min of meer los van dit verschil worden geïnterpreteerd hetgeen opnieuw een geruststelling inhoudt.

### **C** *Correlatie tussen tijd en de variabelewaarden*

Hoewel dezelfde beperkingen, die reeds werden besproken bij het maken van de plots, eveneens golden voor de beoordeelbaarheid van berekende correlatiecoëfficiënten, kunnen deze wel een indruk geven over mogelijke samenhang en de intensiteit daarvan tussen twee variabelen. Significantietoetsing is echter weinig zinvol.

Relevant zijn vooral de correlatiecoëfficiënten tussen tijd (87) en die variabelen, die bij de vergelijking van de totale PO en NO groepen een belangrijke rol zullen spelen. Ik beperk me dan ook tot vermelding van de volgende coëfficiënten: (gegeven in duizendsten).

Hoewel uit deze coëfficiënten niet blijkt dat invloed van de tijd uitgesloten is, mag wel geconcludeerd worden, dat ze allen laag tot zeer laag zijn. Hoogstens lijkt er bij

Tabel 13. Correlatiecoëfficiënten tussen var. 87 en de volgende variabelen.

	Variabele nummer	PO groep	NO groep
Toetsingscriteria	96	— 47	27
	97	36	6
	175	—148	— 42
Weke delen	98	118	68
	101	—185	112
	105	40	74
	109	34	— 61
	111	248	6
	115	171	— 15
	118	76	66
	119	278	112
	124	207	— 33
	125	— 6	— 39
Röntgen gegevens	126	0	71
	130	142	— 70
	131	129	20
	136	— 1	—131
	144	99	— 92
	145	14	103
	146	97	76
	149	98	119
	150	— 65	138
	151	118	22
	152	207	— 38
	153	— 41	— 17

de weke delen gegevens uit de PO groep een lichte tendens tot positieve correlatie met var. 87 te bestaan, hetgeen zou zeggen dat de weke delen verhoudingen in deze groep in de loop der tijd ongunstiger zijn geworden.

Deze tendens is echter in de NO groep afwezig en is dus niet van belang bij vergelijkingen tussen de PO en NO groep, hoogstens zou men kunnen stellen, dat bij een groter interval tussen operatie en naonderzoek bij de PO groep, zodanig dat het identiek wordt aan dat van de NO groep, de weke delenverhoudingen van de PO-groep ongunstiger zouden uitvallen.

Het belangrijkste is echter dat de tijdsafhankelijkheid van de toetsingscriteria in beide groepen verwaarloosbaar lijkt te zijn. Het vergelijken van de PO en NO groepen op basis van de verdelingen van deze criteria lijkt dus, ondanks het grote tijdsverschil, alleszins verantwoord te zijn.

## Hoofdstuk VII

### BEWERKING EN CONCLUSIES

#### A *Inleiding*

Alvorens over te gaan tot de eigenlijke, op het primaire doel van het onderzoek gerichte, gegevensbewerking, is het van belang nog eens te formuleren wat dat doel was. In de inleiding werd reeds beschreven, dat zou worden nagegaan, of het consequent toepassen van de behandelingstechniek volgens Weber tot verbetering van het resultaat zou leiden bij patiënten met een enkelfractuur. Na de bespreking van de toetsingskriteria en de vergelijkbaarheid van de onderzochte patiënten dienen een aantal restricties omtrent de haalbaarheid van het oorspronkelijke doel aangebracht te worden en kunnen we de volgende formulering hanteren:

##### a) Onderzoekdoel

Twee groepen patiënten dienen vergeleken te worden.

- Eén groep bestaat uit patiënten, die behandeld werden, wegens een enkelfractuur, volgens de aanwijzingen die door Weber werden opgesteld (PO groep).
- De andere groep bestaat uit patiënten, die in het gemeenteziekenhuis van Den Haag wegens een enkelfractuur werden behandeld op de in de jaren '62, '63 en '64 gebruikelijke, in hoofdzaak conservatieve, wijze (NO groep).

Ter vergelijking dienen gegevens die door middel van naonderzoek werden verkregen. Deze gegevens worden samengevat in twee gekwantificeerde toetsingskriteria:

- De variabele 97, die als somscore het subjectieve welbevinden ten tijde van het naonderzoek voorstelt en
- De variabele 175, die als somscore de objectieve bevindingen van het naonderzoek samenvat.

Aangezien Weber een duidelijk prognostisch verschil tussen de A-, B- en C-fracturen

meent vast te stellen, dat gunstig beïnvloed zou worden door het toepassen van de door hem geformuleerde behandelingsprincipes, zullen de bovengenoemde vergelijkingen tevens worden uitgevoerd ten aanzien van uit de PO en NO groepen samengestelde subgroepen, die respectievelijk bestaan uit A-, B- en C-fracturen.

#### b) Restricties

- Eventuele significante verschillen, die bij vergelijken van de PO en NO groepen of van subgroepen hieruit, naar voren zullen komen, kunnen behoudens enkele vergelijkbaarheidsrestricties, hoogstens als aangetoond worden beschouwd ten opzichte van de gebruikelijke behandeling van enkelfracturen in het Haagse gemeenteziekenhuis uit de jaren '62, '63 en '64. Iedere generalisatie ten opzichte van conservatieve behandeling van enkelfracturen in het algemeen is, hoewel mogelijk plausibel, altijd hypothetisch.
- De kwantificering van het resultaat stuit op onoverkomelijke moeilijkheden binnen dit onderzoek. Terwijl de subjectieve somscore (97) een redelijke afspiegeling geeft van het subjectief welbevinden van de onderzochte patiënten, blijft de prognostische waarde van de objectieve somscore (175) onzeker. In het vorige hoofdstuk kwam reeds de relatieve onafhankelijkheid van zowel score 97 als 175 ten aanzien van de tijd ter sprake. Dit ondersteunt de vergelijkbaarheid van de PO en NO groepen, al is het niet uitgesloten dat het subjectieve welbevinden van patiënten met een hoge (=ongunstige) objectieve score (175) nog geleidelijk aan zal achteruit gaan.

Op grond van deze situatie moet worden besloten, dat het oorspronkelijk doel, na te gaan of de behandelingstechniek volgens Weber, consequent toegepast bij patiënten met een enkelfractuur, tot resultaatverbetering leidt, niet haalbaar is. Met dien verstande dat een geobjectiveerde meting van het resultaat niet mogelijk is, dat niet exact kan worden bepaald wat een 'beter resultaat' is en dat daarom iedere uitspraak een zekere vrijheid tot subjectieve interpretatie toelaat.

- Bij het bespreken van de vergelijkbaarheid (Hoofdstuk V) bleek, dat een aantal variabelen, die de toetsingscriteria ongunstig beïnvloeden, met een significant hogere frequentie binnen de PO groep voorkwamen dan in de NO groep. Ongunstige bevindingen binnen de PO groep of binnen subgroepen daaruit mogen dus niet zonder meer als een aangetoond gevolg van de operatieve behandeling volgens Weber geïnterpreteerd worden. In de interpretatie van een dergelijke bevinding moeten deze ongunstige verhoudingen binnen de PO groep als mogelijke oorzaak worden gezien.
- In hoofdstuk VI werd de invloed van de naonderzoekperiode op het gemeten behandelingsresultaat besproken. Het bleek dat er weliswaar geen duidelijke invloed van de tijd op de waarde van de toetsingscriteria was aan te tonen, doch dat deze invloed wel op de waarde van enkele andere variabelen leek te bestaan.



Voor het interpreteren van eventuele verschillen, die kunnen bestaan tussen de PO en NO groep ten aanzien van deze variabelen, heeft dat de volgende consequentie:

Een verdelingsverschil tussen de PO en NO groepen van een variabele die aantoonbaar met de naonderzoekperiode samenhangt, kan alleen met het behandelingsverschil tussen PO en NO in verband worden gebracht indien de richting van het verdelingsverschil zodanig is, dat verklaren van het verdelingsverschil uit het verschil in naonderzoekperiode niet mogelijk is.

**B** Vergelijking ten aanzien van de toetsingscriteria, die het behandelingsverschil samenvatten (variabelen 97 en 175). Op basis van het te verwachten resultaat van de operatieve behandeling wordt steeds éénzijdig getoetst. Bij de volgende vergelijkingen wordt de Mann Whitney U-toets gebruikt.

Totale groepen PO en NO

Var. nr.	Toetsingsuitkomst
97	$z=+2,34$ $0,001 < P < 0,01 *$
175	$z=+2,08$ $0,01 < P < 0,02 *$

Dit betekent, dat de groep *geopereerde* patiënten (PO) bij naonderzoek significant BETER uitviel dan de niet geopereerde groep (NO) ten aanzien van het SUBJECTIEF WELBEVINDEN van de patiënten.

De *geopereerde* groep patiënten viel eveneens BETER uit dan de niet geopereerde groep ten aanzien van de SOMSCORE VAN DE OBJECTIEVE BEVINDINGEN bij het onderzoek (175).

Vroegere bevindingen (hoofdstuk VI) waren reeds aanleiding tot het formuleren van aparte somscores voor de niet röntgenologische objectieve gegevens (var. 178) en de röntgenologische onderzoeksgegevens (var. 179); ook op deze criteria werd het totale onderzoekmateriaal getoetst:

Var. nr.	Toetsingsuitkomst
178	$z=-2,72$ $0,001 < P < 0,01 *$
179	$z=+2,84$ $0,001 < P < 0,01 *$

\* Indien P kleiner is dan de significantiegrens, wordt dit met een \* aangegeven.

Dit betekent, dat de *geopereerde groep* patiënten bij naonderzoek significant SLECHTER uitviel dan de niet geopereerde groep patiënten, ten aanzien van de SOMSCORE VAN DE NIET RÖNTGENOLOGISCHE OBJECTIEVE bevindingen.

De *geopereerde groep* patiënten viel echter BETER uit ten aanzien van de SOMSCORE VAN DE RÖNTGENOLOGISCHE bevindingen.

Wij meenden reeds in vorige hoofdstukken een tendens te bespeuren dat de weke delen verhoudingen bij de naonderzochte patiënten zich relatief onafhankelijk ten aanzien van röntgenbevindingen gedroegen. Het lijkt er nu op dat deze tendens ook kon worden aangetoond. Hierbij moeten de interpretatierestricties uit het begin van dit hoofdstuk echter niet uit het oog worden verloren! Hoe deze bevindingen tot een nuancering van de operatieindicaties moeten leiden zal later nog worden besproken.

### *Subgroepen*

Er werd getracht door het vormen van subgroepen uit de totale PO en NO groepen een oordeel te vormen over het effect van de operatieve behandeling ten aanzien van de diverse typen van enkelfracturen. Voor deze typen verwijs ik naar de indeling van Weber zoals deze in hoofdstuk I is besproken. Bij het vormen van deze subgroepen bleken de frequenties vaak zo klein te worden, dat een zinvolle beoordeling onmogelijk was. Verreweg het vaakst komen fracturen van het type I, luxatiefracturen, uit de indeling van Weber voor met als onderindeling de typen A, B en C. Alle beoordelingen zullen dan ook steeds op deze fracturen betrekking hebben:

#### *A-fracturen*

De subgroepen uit PO en NO werden getoetst ten aanzien van de criteria 97, 175, 178 en 179. In geen geval kon een significant verschil worden gevonden.

#### *B-fracturen*

De subgroepen bestaan uit patiënten die primair operatief (PO) of niet operatief (NO) werden behandeld wegens een luxatiefractuur van het bovenste spronggewricht van het B-type:

Var. nr.	Toetsingsuitkomst
97	$z=+2,79$ $0,001 < P < 0,01 *$
175	$z=+2,45$ $0,001 < P < 0,01 *$
178	$z=-1,98$ $0,01 < P < 0,05 *$
179	$z=+4,29$ $P < 0,00003 *$

Dit betekent:

Ten aanzien van het SUBJECTIEF WELBEVINDEN is het resultaat van de *geopereerde* patiënten BETER dan dat bij de niet geopereerde groep.

De SOMSCORE VAN DE OBJECTIEVE BEVINDINGEN is bij de *geopereerde* patiënten GUNSTIGER dan bij de NO groep.

De SOMSCORE VAN DE RÖNTGENBEVINDINGEN is bij de *geopereerde* patiënten GUNSTIGER dan bij de niet geopereerde groep.

De SOMSCORE VAN DE NIET RÖNTGENOLOGISCHE OBJECTIEVE AFWIJ-  
KINGEN is bij de *geopereerde* patiënten echter ONGUNSTIGER dan bij de niet  
geopereerde groep.

### **C-fracturen**

De subgroepen PO en NO bestaan nu uit patiënten, die werden behandeld wegens een luxatiefractuur van het bovenste spronggewricht van het C-type.

Var. nr.	Toetsingsuitkomst
97	$z=+1,43$ $0,5 < P < 0,1$
175	$z=+2,09$ $0,01 < P < 0,05 *$
178	$z=+0,84$ $0,2 < P < 0,3$
179	$z=+2,64$ $0,001 < P < 0,01 *$

Dit betekent:

Ten aanzien van het SUBJECTIEF WELBEVINDEN van de naonderzochte patiënten kon GEEN VERSCHIL worden aangetoond tussen de geopereerde groep en de niet geopereerde groep.

Ten aanzien van de SOMSCORE VAN DE OBJECTIEVE BEVINDINGEN bleek de *geopereerde* groep significant BETER dan de niet geopereerde groep.

\* P is kleiner dan de significantiegrens.

Ten aanzien van de *NIET* RÖNTGENOLOGISCHE OBJECTIEVE BEVINDINGEN kon GEEN VERSCHIL worden aangetoond tussen de geopereerde groep en de niet geopereerde groep.

Ten aanzien van de SOMSCORE VAN DE RÖNTGENOLOGISCHE BEVINDINGEN bleek de *geopereerde* groep significant BETER dan de niet geopereerde groep.

*Vergelijking A-, B- en C-fracturen onderling*

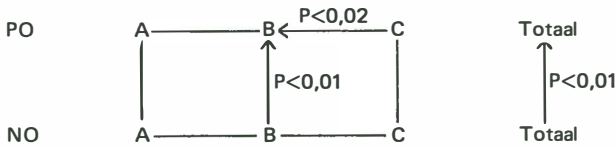
Weber stelt, dat de prognose van een type A-fractuur, bij niet operatieve behandeling gunstiger is dan die van een B-fractuur. Terwijl de B-fractuur op zijn beurt weer gunstiger zou zijn dan een C-fractuur (zie ook hoofdstuk II). Om dit nader na te gaan werden de subgroepen onderling getoetst ten aanzien van de criteria 97, 175, 178 en 179:

A-versus B-fracturen

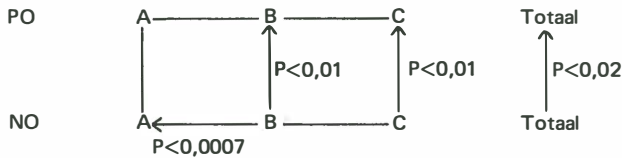
	PO	NO
Var. nr.	Uitkomst	Uitkomst
97	$z=+0,11$ $0,3 < P < 0,5$	$z=+0,88$ $0,1 < P < 0,2$
175	$z=+0,65$ $0,2 < P < 0,3$	$z=+3,29$ $0,0005 < P < 0,0007 *$
178	$z=+1,09$ $0,1 < P < 0,2$	$z=+0,79$ $0,2 < P < 0,3$
179	$z=-0,31$ $0,3 < P < 0,5$	$z=+3,94$ $0,00003 < P < 0,00005 *$
B- versus C-fracturen		
97	$z=+2,14$ $0,01 < P < 0,02 *$	$z=+1,18$ $0,1 < P < 0,2$
175	$z=+1,50$ $0,05 < P < 0,1$	$z=+1,10$ $0,1 < P < 0,2$
178	$z=-0,28$ $0,3 < P < 0,5$	$z=+2,19$ $0,01 < P < 0,02 *$
179	$z=+1,91$ $0,02 < P < 0,05 *$	$z=+0,82$ $0,2 < P < 0,3$

In het voorgaande werden acht steekproeven met elkaar vergeleken, namelijk de totale groepen PO en NO, alsmede de subgroepen A-, B- en C-fracturen uit PO en NO. De bevindingen kunnen in de volgende figuren worden samengevat. De steekproeven worden aangegeven door A, B, C en totaal. Indien het verschil in verdeling van de betrokken variabele tussen twee steekproeven werd getoetst dan is dat aangegeven met een streep of een pijl. Een pijl geeft aan dat een significant verschil werd gevonden, de richting van de pijl wijst naar de steekproef waar de betrokken variabele ten aanzien van het behandelingsresultaat het gunstigst verdeeld was. Tevens is het significantieniveau aangegeven.

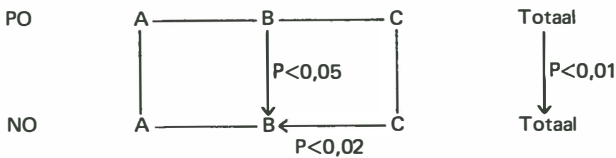
Variabele 97 (somscore subjectief welbevinden)



Variabele 175 (somscore objectieve afwijkingen)



Variabele 178 (somscore weke delen afwijkingen)



Variabele 179 (somscore röntgenologische afwijkingen)

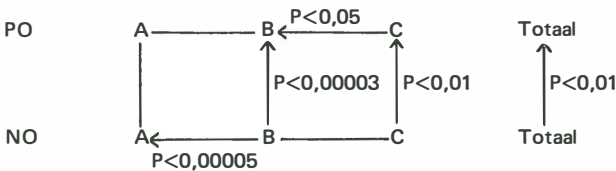


Fig. 12. Verklaring in de tekst.

Uit de toetsingsresultaten kunnen we het volgende concluderen:

Het door Weber veronderstelde prognoseverschil tussen A- en B-fracturen bij niet operatieve behandeling kon inderdaad worden aangetoond voor zover het het röntgenologisch resultaat betreft. De somscore van de niet röntgenologische gegevens bleek bij de A-fracturen en de B-fracturen ongeveer gelijk uit te vallen. Ook de prognose aangaande het subjectieve welbevinden van de patiënten bleek bij de patiënten met een A-fractuur niet te verschillen van die bij patiënten met een B-fractuur.

Het consequent opereren van patiënten met een A-fractuur of een B-fractuur doet het prognostisch verschil ten aanzien van het röntgenologisch resultaat wegvallen zodat bij de geopereerde patiënten geen enkel prognostisch verschil tussen patiënten met een A-, en patiënten met een B-fractuur kon worden aangetoond.

Het prognostisch verschil tussen patiënten met een B-fractuur en patiënten met een

C-fractuur is zowel binnen de PO groep als de NO groep aantoonbaar. Met dien verstande, dat er binnen de PO groep voor de patiënten met een C-fractuur een duidelijk slechtere prognose bestaat ten aanzien van het subjectief welbevinden en het röntgenologisch resultaat, terwijl er op deze punten bij de NO groep geen duidelijk verschil aangetoond kon worden. Bij deze laatste groep blijkt juist de verhouding van de weke delen ongunstiger te zijn bij de C-fracturen.

Het lijkt voor de hand liggend, dat het verschil in röntgenologisch resultaat tussen de B- en C- subgroepen van de geopereerde patiënten ook het prognoseverschil in subjectief welbevinden kan verklaren. Zoals reeds vroeger ter sprake is geweest hangen deze twee gegevens echter niet zodanig samen, dat deze verklaring zonder meer aanvaard mag worden. Trouwens het sterke verschil tussen de röntgenologische prognose van patiënten met een B-fractuur en patiënten met een A-fractuur ging niet samen met een aantoonbaar verschil ten aanzien van het subjectief welbevinden. Waarschijnlijker is dat de prognose van de B-fracturen door de operatieve behandeling sterker ten gunste wordt beïnvloed dan die van de C-fracturen, zodat een aantoonbaar verschil in het subjectieve welbevinden gaat ontstaan, dat min of meer los staat van de röntgenologische bevindingen.

Deze opvatting kan steun vinden door het feit dat binnen de groep C-fracturen de röntgenologische bevindingen door de operatie wel sterk lijken te zijn verbeterd (zie de tabel op pag.104), doch dat het subjectieve welbevinden (97) bij de patiënten met een geopereerde C-fractuur niet significant is verbeterd ten opzichte van dat van de vergelijkbare patiënten, die niet werden geopereerd. De verbetering in het röntgenaspect heeft dus geen duidelijk gevolg gehad in een verandering in het subjectief welbevinden.

Een andere benadering van de B-C verschillen binnen de PO groep is de veronderstelling, dat zich binnen de C-fracturen meer sterk gedislodeerde of sterk communiteieve fracturen kunnen bevinden, factoren die reeds bleken met toetsingscriteria samen te hangen.

Zou hier de verklaring gezocht moeten worden, dan zouden de gevonden B-C verschillen dus niet uit verschillende fractuurtypes, doch uit de oorspronkelijke dislocatie, met de daaraan gepaard gaande beschadigingen, of uit het communiteieve karakter verklaard moeten worden. De oorspronkelijke verklaring van Weber uit het verscheurd zijn van de voorste syndesmose zou dan op de achtergrond raken. Om dit nader te analyseren zouden subgroepen naar dislocatiegraad moeten worden gemaakt, die weer worden onderverdeeld naar type fractuur. Na deze sterke onderverdeling bleven echter dermate kleine groepen over, dat beoordeling niet meer zinvol was en het materiaal bleek te klein om een uitspraak over dit probleem te kunnen geven.

Tenslotte kan worden gesteld, dat er inderdaad prognostische verschillen tussen A-, B- en C-fracturen konden worden vastgesteld, dat de verschillen tussen de A- en B-fracturen door de operatieve therapie lijken te zijn verminderd, dat de verschillen

tussen de B- en C-fracturen zowel in de geopereerde groep, als de niet geopereerde groep aantoonbaar waren doch dat deze zich wel openbaarden ten opzichte van verschillende toetsingscriteria, en dat een slechtere prognose ten aanzien van het röntgenologisch resultaat niet altijd samengaat met een slechtere prognose van het subjectieve welbevinden of omgekeerd.

### C Nadere analyse van het gedrag van de somscores 175, 178 en 179.

Hierbij is het van belang er op te wijzen dat:

- 175 de somscore van alle fysisch diagnostische en röntgenologische gegevens vertegenwoordigt, dat
- 178 de somscore is van alle fysisch diagnostische naonderzoekgegevens alleen en als zodanig vooral de toestand van de weke delen vertegenwoordigt en dat
- 179 de somscore is van alle röntgenologische gegevens en het anatomisch herstel en de toestand van de benige gewrichtselementen vertegenwoordigt.

Op meerdere plaatsen wees ik er reeds op dat de subjectieve somscore (97) zich t.o.v. de gegevens van de weke delen en de anatomische gegevens van het skelet verschillend leek te gedragen, terwijl de relatie tussen de weke delen en de anatomische röntgegegevens zich niet zodanig opstelde dat deze afwijkingen onverbrekelijk samen lijken te gaan. Het lijkt nuttig de aanleidingen tot deze indruk nog eens bij elkaar te zetten. Aangezien het bij het behandelen van patiënten uiteindelijk gaat om hun subjectief welbevinden zal variabele 97 bij deze beschouwingen een centrale rol spelen.

Aangezien het bij het behandelen van patiënten echter ook gaat om het handhaven van het subjectieve welbevinden, m.a.w. om het bereiken van een goede prognose ten aanzien van het subjectief welbevinden, zal ook het gedrag in de tijd van de diverse scores in de overwegingen worden betrokken.

#### a) De tijd

Om enig inzicht in de prognose te krijgen werd nagegaan of toenemen van de naonderzoekperiode van invloed was op de waarde van de toetsingscriteria. Daartoe werden rangcorrelaties berekend met deze periode (87).

Ze werden in de volgende tabel opgenomen:

Tabel 14. Rangcorrelatiecoëfficiënten tussen var. 87 enerzijds en de variabelen 97, 175, 178 en 179 anderzijds. De coëfficiënten worden gegeven in duizendsten.

Var. nr	PO	NO
87— 97	+28	+16
87—175	+241*	+ 3
87—178	+357*	—70
87—179	—3	+35

\* P is kleiner dan de significantiegrens.

Hieruit blijkt, dat alleen de weke delen verhoudingen, voor zover ze worden uitgedrukt in de score 178, binnen de PO groep duidelijk met de grootte van de naonderzoekperiode samenhangen. D.w.z. dat moet worden verwacht, dat de weke delen verhoudingen na de primaire genezing geleidelijk verslechteren.

De onafhankelijkheid van 178 ten opzichte van 87 binnen de NO groep wijst erop, dat de weke delen verhoudingen zich na ong. vijf jaar (de kortste naonderzoekperiode binnen de NO groep) stabiliseren, of de verslechtering zo langzaam gaat verlopen, dat de afhankelijkheid binnen dit onderzoek niet meer kon worden vastgesteld.

Het subjectief welbevinden lijkt zich in beide groepen relatief onafhankelijk van de tijd op te stellen, evenals de röntgenologische gegevens. Nu is het slecht denkbaar, dat steeds slechter wordende weke delen op den duur het subjectief welbevinden ongemoeid zullen laten, hypothetisch kan dit leiden tot de opvatting, dat het subjectief welbevinden langzamer achteruit gaat dan de weke delen verhoudingen zodat binnen de tijdspreiding van de binnen dit onderzoek onderzochte patiënten deze geleidelijke achteruitgang (nog) niet aantoonbaar was.

De relatieve onafhankelijkheid ten aanzien van de tijd bij de röntgengegevens is moeilijker te verklaren; vier hypothesen zijn mogelijk:

- een gemeenschappelijke factor (b.v. posttraumatische ischaemie) doet in een betrekkelijk korte tijd de toestand van de weke delen achteruit gaan, doch heeft meestal minder en langzamer effect op de benige gewrichtsdelen. In dit geval zouden dus ook de röntgenaspecten op den duur kunnen verslechteren, maar die hebben als zodanig geen direct causaal verband met de objectieve achteruitgang van de weke delen.
- Een zeer langzaam proces in de benige gewrichtsdelen doet een reactie in de omgeving ontstaan, die tot een relatief snelle achteruitgang van de weke delen leidt. In dit geval zou de oorzaak van de weke delen afwijkingen wel in de botafwijkingen moeten worden gezocht.
- De situatie van de weke delen en die van de benige gewrichtselementen staan geheel los van elkaar en is alleen gebonden door het gemeenschappelijk letsel uit het verleden. De weke delen gaan nog geleidelijk achteruit en de benige elementen zijn in een eindtoestand, die onafhankelijk is van de tijd.
- De afwijkingen van de weke delen hebben de fysiologische verhoudingen van het gewricht zodanig verstoord, dat op den duur ook de benige gewrichtselementen zodanig achteruitgaan, dat degeneratieve kenmerken ook op de röntgenfoto zichtbaar worden. Deze ontwikkelen zich echter op grond van en daarom ook later dan de zichtbare weke delen afwijkingen. Nu moet de causale factor in de weke delen worden gezocht en is eenzelfde situatie geschetst als bij het subjectief welbevinden werd verondersteld: de röntgenologische botafwijkingen ontwikkelen zich zo langzaam dat hun tijdafhankelijkheid binnen de tijdslijmieten van dit onderzoek niet kon worden aangetoond.



Hoe het ook zij, ten aanzien van het gedrag in de tijd stellen de weke delen verhoudingen zich duidelijk anders op dan het subjectief welbevinden of de waar te nemen röntgenafwijkingen.

#### b) De factoranalyse

In hoofdstuk V werd een factoranalyse besproken van de samenstellende variabelen van somscore 175 (Zie pag. 90). De eerste factor bleek uitsluitend röntgenologische gegevens van de benige gewrichtsdelen te verzamelen. In factor 2 en 3 kwamen andere gegevens samen.

Ten aanzien van de factoranalyse van alle objectieve gegevens gedragen de weke delen gegevens zich dus anders als de röntgenologische gegevens.

#### c) De vergelijking van de subgroepen

Bij de vergelijking van de PO en NO groepen in dit hoofdstuk (VII B) bleek de PO groep gunstiger dan de NO groep ten aanzien van de somscores 97 en 179, terwijl ten opzichte van de score 178 juist de NO groep gunstiger uitviel. Bij vergelijking van de B-fracturen bleek eveneens de score 178 in tegengestelde richting te verschillen van die van de scores 97 en 179. Men zou nu kunnen veronderstellen, dat de subjectieve score meer dan de score 179 is gebonden dan aan de weke delen verhoudingen (178). Dit blijkt echter niet bevestigd te worden bij de vergelijking van de C-fracturen waarbij geen verschil in subjectief welbevinden aantoonbaar was (97), terwijl de weke delen verhoudingen van de geopereerde patiënten en de niet geopereerde patiënten ook niet duidelijk verschilden. Bij deze C-fracturen was echter wel een duidelijke verbetering bij de PO groep ten aanzien van de somscore van de röntgenologische gegevens (179).

Ook bij de vergelijkingen van de diverse fractuurtypen onderling bleken de anatomische verhoudingen blijkens de röntgenfoto sterk te kunnen verschillen, terwijl er subjectief geen verschil werd gevonden (B.v. A-versus B-fracturen uit de NO groep). Bij B- versus C-fracturen uit de PO groep blijken verbeteringen blijkens de röntgenfoto juist wel samen te gaan met een subjectieve verbetering, terwijl bij dezelfde vergelijking binnen de NO groep het ook mogelijk blijkt dat alleen de weke delen verhoudingen verschillen zonder dat dit gepaard gaat met aantoonbare verschillen in het subjectief welbevinden (97) of de röntgenologische bevindingen (179).

Dit alles wijst er op dat de drie scores weliswaar verwant zijn, doch dat zij toch ook zo los samenhangen, dat verschillen in onderling tegengestelde richtingen kunnen ontstaan, als de omstandigheden dat meebrengen. Het grillige gedrag van de somscores onderling, doet veronderstellen dat ze of geheel niet samenhangen of dat er buiten de nu gebruikte indeling in A-, B- en C-fracturen een veel belangrijker factor

een rol speelt in het bepalen van de prognose. Deze factor zou dan in bepaalde gevallen een zo belangrijke invloed kunnen hebben, dat verbeteren van de anatomische verhoudingen van het skelet niet meer bijdraagt aan het verbeteren van de prognose ten aanzien van het subjectief welbevinden van de patiënt.

d) De onderlinge samenhang van de somscores.

De veronderstelling, dat de somscores in het geheel niet zouden samenhangen werd nagegaan aan de hand van correlaties tussen de somscores onderling.

Omdat het bij het afwegen van het behandelingsresultaat bijzonder belangrijk lijkt te zijn of de patiënt zelf tevreden is met dat resultaat werd ter vergelijking de variabele die dat weergeeft (96) in de tabellen opgenomen:

Bij de PO en NO groep werden de volgende correlatiecoëfficiënten gevonden (weergegeven in duizendsten).

Tabel 15. Verklaring in de tekst.

Var. nr.					Var. nr.				
PO (N=85)					NO (N=156)				
96	—				96	—			
97	779	—			97	852	—		
175	592	540	—		175	298	436	—	
178	573	594	737	—	178	364	515	720	—
179	456	347	902	407	179	253	364	959	512
<hr/>					<hr/>				
Var. nr.	96	97	175	178	Var. nr.	96	97	175	178

Hoe moet dit alles worden geïnterpreteerd?

In de eerste plaats blijkt de objectieve totaalscore (175) in beide groepen sterker met 179 dan met 178 samen te hangen. Dat klopt aangezien er aan 175 een groter aantal röntgegegevens dan weke delen gegevens bijdragen, terwijl alle samenstellende variabelen gelijk deelnemen: als een samenstellende variabele ≠ 0 is dan draagt hij met 1 punt aan de somscore bij. Het is echter van belang te noteren dat 175 en 179 zo sterk samenhangen (met een correlatiecoëfficiënt van >900) dat het onwaarschijnlijk lijkt dat ze zich bij bewerkingen verschillend zullen gedragen. Deze verwachting werd echter bij de bewerkingen reeds éénmaal gelogenstraft bij de vergelijking tussen de 8- en C-fracturen uit de PO groep waar het verschil in score 179 de significantiegrens wel en dat in score 175 deze grens niet bleek te halen. De samenhang tussen beide scores laat blijkbaar genoeg vrijheid om aan de weke delen verhoudingen (178) de gelegenheid te geven een absoluut samengaan te verhinderen. In de tweede plaats blijken alle in de tabellen opgenomen scores samen te hangen, dit was uiteraard te verwachten, doch de veronderstelling, dat de somscores niet zouden samenhangen, als verklaring voor hun grillig onderling gedrag, gaat dus niet op.

Ten derde blijkt vooral in de PO groep de samenhang tussen 178 en 179 onderling veel geringer te zijn dan met hun somscore (175). Hoewel dit vaak met somscores

en hun samenstellende variabelen het geval is, verklaart dit tevens, dat hun samenhang los genoeg is om zich bij toetsing van verdelingsverschillen soms tegenovergesteld op te kunnen stellen.

Tenslotte blijkt zowel de tevredenheid van de patiënt zelf (96) als de somscore van het subjectief welbevinden (97) in beide groepen nauwer met de weke delen verhoudingen (178) samen te hangen, dan met de waargenomen afwijkingen op de röntgenfoto (179). Dit is van groot belang aangezien hieruit direct volgt, dat voor het welslagen van de behandeling de op de röntgenfoto waargenomen afwijkingen slechts een beperkte betekenis blijken te hebben! De grote invloed die de toestand van de weke delen blijkt te hebben op het subjectief welbevinden van de patiënten brengt de bevinding, dat juist deze verhoudingen in de PO groep slechter bleken te zijn dan in de NO groep, in een speciaal daglicht. Men moet zich afvragen of de operatieve behandeling van de fractures de weke delen verhoudingen nadelig heeft beïnvloed.

Ik kom hier nog op terug. Thans is het van belang tot de conclusie te komen dat:

- de weke delen verhoudingen relatief matig samenhangen met de röntgenologische toestand van het skelet van het bovenste spronggewricht en
- dat het subjectief welbevinden sterker lijkt samen te hangen met de toestand van de weke delen dan met die van de röntgenologische toestand van de benige elementen van het bovenste spronggewricht.

**D** *Worden weke delen verhoudingen door de operatieve behandeling belangrijk nadelig beïnvloed?*

Om deze vraag te benaderen moet herinnerd worden aan twee reeds beschreven bevindingen:

- a) De score 178 (somscore van de weke delen bevindingen) was in de groep PO duidelijk afhankelijk van de naonderzoekperiode, i.e. het tijdsverloop tussen ongeval en naonderzoek.
- b) De verdeling van 178 bleek binnen de groep PO ongunstiger te zijn dan bij de NO groep.

Thans moet worden verwezen naar de derde interpretatierestrictie uit het begin van dit hoofdstuk: het interpreteren van voor PO ongunstig uitvallende bevindingen dient met de grootste voorzichtigheid te gebeuren, aangezien in deze groep een aantal variabelen in een ongunstige verdeling voorkomen ten opzichte van de NO groep.

Met deze restrictie voor ogen kunnen de in PO ongunstig verdeelde variabelen worden bekeken teneinde voor de ongunstige weke delen verhoudingen van de PO groep in de verdelingsverschillen tussen PO en NO een verklaring te vinden, die dan onafhankelijk van de operatieve therapie zou zijn.

Het betreft de variabelen 5, 6, 12, 13, 14, 38 en 67.

Van deze variabelen is het alleen van variabele 67 zonder meer duidelijk dat van een ongunstige verdeling van deze variabele (de primaire dislocatie) een ongunstige weke delen situatie het gevolg kan zijn.

Het effect van deze variabele op het resultaatverschil tussen PO en NO kan uitgeschakeld worden door steeds groepen met gelijke dislocatiegraad te vergelijken. Op deze wijze werd de volgende tabel samengesteld:

Tabel 16 Toetsing van verdelingsverschillen tussen subgroepen uit PO en NO, die steeds een gelijke dislocatiegraad hebben. Verdelingen worden beoordeeld van variabelen 96, 97, 175, 178 en 179. Steeds tweezijdig getoetst.

	67=1	67=2	67=3	67=4	67=5**
Aantallen					
PO	14	22	23	8	8
NO	31	45	24	13	2
Varia- bele 96	$z=+1,75$ $0,05 < P < 0,1$	$z=+2,01$ $0,02 < P < 0,05^*$	$z=+0,35$ $0,5 < P < 0,7$	$z=+1,31$ $0,1 < P < 0,2$	$U=1,0$ $0,05 < P < 0,1$
Varia- bele 97	$z=+1,83$ $0,05 < P < 0,1$	$z=+2,27$ $0,02 < P < 0,05^*$	$z=+0,99$ $0,3 < P < 0,5$	$z=+0,80$ $0,3 < P < 0,5$	$U=1,0$ $0,05 < P < 0,1$
Varia- bele 175	$z=+0,047$ $0,95 < P < 0,98$	$z=+2,47$ $0,01 < P < 0,02^*$	$z=+3,50$ $0,0005 < P < 0,0007^*$	$z=+1,50$ $0,1 < P < 0,2$	$U=4,5$ $0,5 < P < 0,7$
Varia- bele 178	$z=-0,45$ $0,5 < P < 0,7$	$z=-0,45$ $0,5 < P < 0,7$	$z=+0,62$ $0,5 < P < 0,7$	$z=-0,59$ $0,5 < P < 0,7$	$U=6,0$ $0,7 < P < 0,8$
Varia- bele 179	$z=+1,01$ $0,3 < P < 0,5$	$z=+3,37$ $0,0003 < P < 0,0005^*$	$z=+4,10$ $0,00005 < P < 0,00007^*$	$z=+1,90$ $0,05 < P < 0,1$	$U=6,5$ $0,8 < P < 0,9$

\*\* Exacte toets wegens kleine getallen.

Uit deze tabel blijkt duidelijk dat het grote verschil in de toestand van de weke delen, dat blijktens de significante verschillen in de score 178 steeds tussen de groepen PO en NO leek te bestaan, geheel verdwenen is. Hoewel soms significante verschillen in het subjectief welbevinden (97) of de röntgenbevindingen (179) bestaan, beweegt de overschrijdingskans van score 178 zich steeds rond de 0,5.

Nu zou men geneigd zijn te concluderen, dat de bij het naonderzoek waargenomen toestand van de weke delen geheel toegeschreven kan worden aan de oorspronkelijke dislocatie en dat het al of niet opereren daar geen invloed op heeft gehad. Dan dient men zich echter de vierde restrictie uit het begin van dit hoofdstuk te herinneren en zich te realiseren, dat juist de variabele 178 binnen de groep PO duidelijk samenhang met de grootte van het tijdsinterval tussen fractuur en naonderzoek. Bij een oorspronkelijk gelijke en ondanks de operatie gelijk blijvende weke delen beschadiging, zou men verwachten, dat binnen de groep NO, door het veel langer

\* P is kleiner dan de significantiegrens.

gemiddeld tijdsinterval tussen letsel en naonderzoek, ongunstiger weke delen verhoudingen zouden zijn gevonden dan bij de vergelijkbare PO groep. Nu dit echter niet het geval is ligt de conclusie voor de hand, dat de operatie een ongunstig effect heeft gehad op de bij het naonderzoek gevonden weke delen toestand, zodat deze in de gemiddeld kortere naonderzoekperiode in dezelfde toestand is gekomen als die van de niet geopereerde groep in een veel langere tijd.

We kunnen hieruit de volgende conclusies trekken:

- De bij het naonderzoek gevonden weke delen afwijkingen worden vrijwel volledig bepaald door de beschadiging die bij het oorspronkelijk letsel, waarbij ook de fractuur ontstond, optrad.
- De operatieve behandeling benadeelt de weke delen verhoudingen rond het gewricht, doch niet zodanig, dat de betere anatomische repositie, die door middel van de operatieve behandeling mogelijk is niet in veel gevallen tot een uiteindelijk beter resultaat zou kunnen leiden.
- De op pag. 109 reeds veronderstelde factor, die belangrijker is voor het bepalen van de prognose dan de indeling van de fractuur in een A-, B- of C-type, blijkt de dislocatie te zijn. Deze blijkt vrijwel bepalend te zijn voor de uiteindelijke weke delen verhoudingen, die op hun beurt sterker samenhangen met het uiteindelijk subjectief welbevinden van de patiënt dan de verhoudingen van het benige skelet, voor zover die op de röntgenfoto te beoordelen zijn.
- De relatief ongunstige invloed van de operatieve behandeling maakt het noodzakelijk na te gaan of een nauwkeuriger operatieindicatie mogelijk is dan die welke door Weber werd geïntroduceerd. Aan dit vraagstuk zal het volgende onderdeel van dit hoofdstuk gewijd zijn.

### ***E Is het mogelijk een optimale operatieindicatie uit de beschikbare naonderzoekgegevens te berekenen?***

Om een optimale operatieindicatie te bepalen dient eerst te worden vastgesteld welke voorwaarden aan het resultaat dienen te worden gesteld. Deze zijn als volgt te formuleren:

- 1) In verband met de sterke verwantschap tussen de weke delen verhoudingen en het subjectief welbevinden van de patiënten alsmede door het geleidelijk slechter worden van de weke delen verhoudingen in de tijd dient score 178 bij het 'beter' resultaat na eenzelfde periode minstens even goed te zijn als bij het 'slechter' resultaat.
- 2) Het 'beter' resultaat dient in gunstige zin significant te verschillen van het 'slechter' resultaat ten aanzien van het subjectief welbevinden van de patiënten (score 97).
- 3) Hoewel de score van de röntgenologisch waargenomen afwijkingen (179) binnen dit onderzoek geen duidelijke relatie met de naonderzoekperiode had, wordt

er voorlopig toch van uitgegaan, dat een gunstiger score op de lange duur met een betere subjectieve prognose samenhangt. Dit zal echter met een beter gericht onderzoek geverifieerd moeten worden. In dit verband dienen voorlopig aan score 179 dezelfde voorwaarden gesteld te worden als boven voor 178 werden beschreven.

Met deze drie voorwaarden voor ogen blijkt dat uit de vergelijking tussen de geopereerde en de niet geopereerde patiënten uit de A-, B- en C- groepen geen conclusie te trekken valt over de 'betere' behandeling. Daar waar score 97 en 179 ten gunste van de geopereerde groep patiënten significant van die bij de niet geopereerde patiënten verschilt, verschilt score 178 juist ten ongunste van de geopereerde patiënten, terwijl bij de C-fracturen geen verschil in score 97 kon worden aangetoond.

We moeten dus vaststellen, dat de indeling van Weber teleurstelt als hulpmiddel om de optimale operatieindicatie vast te stellen. We hebben echter juist vastgesteld, dat score 178 nauw samenhangt met de dislocatie en dat de indeling naar de dislocatiegraad groepen oplevert, die niet verschillen wat hun score 178 betreft. Het blijkt nu mogelijk aan de hand van tabel 16 vast te stellen bij welk type fracturen winst verwacht mag worden bij operatieve therapie.

Bij 67=1 blijkt ten aanzien van geen enkel toetsingskriterium een significant verschil tussen de geopereerde en de niet geopereerde patiënten te kunnen worden aangetoond. Het lijkt erop of het subjectief welbevinden en de tevredenheid van de patiënt tendeert naar een gunstiger situatie bij de geopereerde groep, doch gezien het totaal ontbreken van enige aanwijzing bij de objectieve gegevens lijkt het ontoelaatbaar hier enige betekenis aan te hechten.

Bij 67=2 blijken de verschillen aan alle gestelde voorwaarden te voldoen om tot de uitspraak: de primair operatieve behandeling is 'beter' te kunnen komen. Er zij op gewezen, dat ondanks het grote verschil in het aantal waargenomen afwijkingen op de röntgenfoto (score 179) het verschil in subjectief welbevinden maar juist aan de afgesproken significantiegrens voldoet.

Bij 67=3 blijkt de operatieve behandeling een enorme verbetering te geven ten aanzien van de röntgenbevindingen bij naonderzoek. Het subjectief welbevinden van de patiënten vertoont ondanks deze verbetering echter geen verschil, het tendeert er zelfs niet naar. Hier moeten we concluderen, dat de grote weke delen beschadiging die met het letsel is ontstaan het subjectief resultaat heeft bepaald en dat een conservatieve behandeling in staat is de anatomische skeletrelaties zover te verbeteren, dat afwijkingen van een perfecte repositie, die mogelijk overblijven, een ondergeschikte rol voor het subjectief welbevinden spelen.

Helaas moet dus geconcludeerd worden, dat de röntgenfoto weliswaar voortreffelijk werd verbeterd, doch dat de patiënt daar niet van heeft kunnen profiteren.

Bij 67=4, dus de nog sterker gedислоceerde fracturen, blijken ook bij de geopereerde

patiënten zoveel secundaire röntgenologische afwijkingen te zijn ontstaan, dat zelfs score 179 geen duidelijk verschil meer oplevert tegenover die van de niet geopereerde patiënten. Hier is de primaire beschadiging dus van zowel weke delen als benige gewrichtselementen zodanig, dat operatieve correctie geen meetbare verbetering kon geven op het late resultaat.

Bij 67=5 zijn de groepen zo klein geworden, dat ook het discriminerend vermogen van de toetsing klein wordt. Het niet kunnen aantonen van een verschil (in welke richting dan ook!) toont steeds minder aan dat het er ook niet is. Gezien de bevindingen bij de vorige dislocaties, dient men hier echter zeer terughoudend te zijn met aan te nemen, dat operatieve behandeling voordeel moet bieden.

Uit deze beschouwingen kunnen we dus de conclusie trekken:

De oefenstabiele osteosynthese voor luxatiefracturen van het bovenste spronggewricht is optimaal geïndiceerd voor fracturen, die:

lateralisatie van de talus groter dan 0, doch hoogstens 3 mm en/of meer dan 2 mm doch hoogstens 4 mm verkorting van de fibula en/of niveauverschil van de tertius tot hoogstens 2 mm en/of vorkverwijding tot hoogstens 3 mm.

D.w.z. fracturen uit de groep 67  $\approx$  2.

In het algemeen dislocaties groter dan 2 mm doch hoogstens 4 mm vertonen op de primaire ongevalsfoto, d.w.z. fracturen uit de groep

Van operatieve behandeling van alle andere enkelfracturen uit de groep luxatiefracturen, mag gezien de resultaten van dit onderzoek geen voordeel voor de patiënt worden verwacht ten aanzien van een alternatieve conservatieve behandeling.

Het is hier van belang er nog eens op te wijzen, dat de conclusie, die zojuist werd getrokken, gebaseerd is op het ontbreken van een verschil in score 178 tussen de subgroepen uit PO en NO ingedeeld naar hun dislocatiegraad. Terughouding bij de interpretatie hiervan is noodzakelijk vooral op grond van de tijdafhankelijkheid van score 178 en het grote verschil in gemiddelde naonderzoek periode tussen de groepen PO en NO. Het zou dus kunnen zijn dat bij een onderzoek waarbij het tijdinterval tussen fractuur en naonderzoek constant is, de score 178, i.e. de weke delen verhoudingen, en daarmee het subjectieve welbevinden, ook in groep 67=2 niet meer ten gunste van de geopereerde groep zouden uitvallen.

#### **F Samenvatting van de belangrijkste conclusies uit dit hoofdstuk**

a) Bij het beoordelen van het resultaat van de behandeling van luxatiefracturen van het bovenste spronggewricht blijken voornamelijk twee groepen waarnemingsgegevens op het subjectief welbevinden van de patiënten en op de prognose daarvan invloed te hebben.

Deze twee groepen vertonen onderling slechts een matig sterke samenhang. Als waargenomen afwijking bij het naonderzoek beïnvloeden ze beiden in negatieve zin het subjectief welbevinden van de patiënten. Slechts één van beide groepen is vatbaar voor operatieve positieve beïnvloeding doch dat heeft onverbiddelijk een negatieve beïnvloeding van de andere groep ten gevolge. Het betreft:

- De groep weke delen gegevens, die een wat nauwere samenhang met het subjectief welbevinden heeft dan de tweede groep en tevens het sterkste met het tijdsinterval tussen fractuur en naonderzoek samenhangt en als zodanig voor een belangrijk deel de prognose bepaalt.
- De op de röntgenfoto waargenomen afwijkingen, die weliswaar samenhangen met het subjectief welbevinden van de patiënt, doch minder sterk dan de vorige groep, doch die sterk positief zijn te beïnvloeden door operatieve correctie van de op de röntgenfoto waargenomen afwijkingen.

b) De groep weke delen gegevens blijkt nergens door de operatieve behandeling positief te zijn beïnvloed, er is echter wel een aanwijzing, dat het operatief ingrijpen een negatieve invloed op de bij het naonderzoek waargenomen weke delen afwijkingen heeft.

c) De uiteindelijke toestand van de weke delen blijkt voor een belangrijk deel bepaald te worden door het weke delen letsel dat tegelijk met de fractuur ontstond. Dit is met nauwkeurigheid af te lezen aan de primaire dislocatie op de ongevalsfoto (variabele 67).

De indeling van de patiënten naar de rangorde van variabele 67 bleek tot het wegvallen van de verschillen ten aanzien van de weke delen score tussen de geopereerde en de niet geopereerde patiënten (met de zelfde dislocatiegraad) te leiden.

d) In tegenstelling tot de indeling naar dislocatiegraad, bleek de fractuurindeling volgens Weber niet te kunnen leiden tot een duidelijke indruk van het effect van de operatieve behandeling, aangezien ofwel de weke delen verhoudingen tegengesteld verschilden van andere toetsingscriteria, dan wel geen duidelijk verschil in subjectief welbevinden van de patiënt aantoonbaar was.

e) Door de belangrijke invloed van de weke delen verhoudingen op het subjectief welbevinden van de patiënt en op de prognose daarvan, terwijl ze door operatie alleen negatief beïnvloed kunnen worden, bevinden ze zich in een soort 'op de wip' positie:

- indien de primaire weke delen beschadiging niet al te groot is (af te lezen aan de primaire dislocatie) dan kan een beter herstel van de benige relaties van het gewricht dan door conservatieve behandeling mogelijk is, het uiteindelijk resultaat verbeteren. De extra beschadiging van de weke delen, door de operatie, speelt dan een ondergeschikte rol. Dit betreft de groep fracturen 67=2.
- indien daarentegen de primaire weke delen beschadiging wel groot is blijkt een



grote primaire dislocatie, dan bepaalt deze voor zo'n groot deel het welbevinden van de patiënt, dat het operatief streven naar een betere röntgenologische repositie dan met conservatieve middelen mogelijk is, geen verbetering voor de patiënt kan brengen, doch in tegendeel alleen een ongunstige invloed kan hebben op de weke delen verhoudingen.

f) Het spreekt vanzelf, dat voor deze wat onverwachte bevindingen binnen dit hierop niet direct gerichte onderzoek slechts aanwijzingen, zij het vrij harde, zijn gevonden.

Vooraf door het grote verschil in tijdinterval tussen fractuur en naonderzoek bij de vergeleken groepen, is het niet onwaarschijnlijk dat de weke delen waarnemingen van geopereerde patiënten bij een anders gericht onderzoek ongunstiger zouden uitvallen.

g) Tenslotte dient er op te worden gewezen, dat nader onderzoek noodzakelijk is teneinde de betekenis van de diverse typen van fractuurdetails te leren kennen, doch dan onafhankelijk van de dislocatiegraad. Binnen dit onderzoek was dat niet mogelijk aangezien restdislocaties als b.v. lateralisatie steeds ook samen bleken te gaan met een grotere primaire dislocatie zodat de ongunstige betekenis van lateralisatie alleen al door de grotere primaire dislocatie zou worden verklaard. Het onderverdelen van de subgroepen dislocatiegraad uit tabel 16 leidt tot zulke kleine groepen, dat beoordeling niet zinvol is.

## SAMENVATTING

Het doel van het in dit boek beschreven onderzoek is het nagaan van de betekenis van de door Weber, in zijn monografie over de letsels van het bovenste spronggewricht, gepropageerde oefenstabiele osteosynthese als voorkeursbehandeling van de luxatiefracturen van het bovenste spronggewricht. Resultaten bij een groep van 97 volgens deze methode behandelde patiënten worden vergeleken met resultaten, die werden verkregen met een in principe conservatieve behandeling van deze fracturen. De operatieindicaties en fractuurindeling, die door Weber naar voren worden gebracht berusten op hypothesen, die hij afleidt uit een aantal anatomische en functionele beschouwingen. Daarom worden deze beschouwingen in hoofdstuk I besproken. De belangrijkste conclusies uit deze beschouwingen luiden:

- 1) De talus maakt t.o.v. de tibia geen simpele dorsaal-plantair flexie beweging met één enkele rotatieas, doch een ingewikkelde draaibeweging waarbij drie verschillende draaiingsassen zijn te onderscheiden.
  - 2) Dientengevolge dient de fibula met het draaiende en kantelende laterale zijvlak van de talus mee te bewegen teneinde het contact tussen de gewrichtsvlakken te onderhouden. Vrije bewegingen van de fibula ten opzichte van de tibia zijn dus obligaats voor het ongestoord functioneren van het bovenste spronggewricht.
  - 3) In enkele fases van het lopen blijkt de talus met grote kracht tegen de laterale malleolus aan te dringen.
  - 4) Op grond hiervan hebben de ligamenteuze verbindingen tussen de fibula en tibia juist tijdens de genoemde loopfases een belangrijke functie om subluxaties van talus ten opzichte van het distale tibia plateau te voorkomen.
- Weber leidt uit deze beschouwingen de volgende hypothesen af:

- 1) Voor het bereiken van een optimale genezingskans bij fracturen van het bovenste spronggewricht is een volledig stabiele enkelvork essentieel.
- 2) De stabiliteit van de enkelvork is ernstig bedreigd bij het insufficiënt zijn van één der ligamenteuze verbindingen tussen fibula en tibia. Van deze verbindingen, de

z.g. vorkbanden of syndesmoses is vooral de voorste bij luxatiefracturen van het bovenste spronggewricht vaak verscheurd.

3) Door de grote krachten, die reeds bij gewoon lopen tussen de verschillende gewrichtsvlakken van het bovenste spronggewricht (B.S.G.) optreden geeft iedere onvolledige gewrichtsvlakcongruentie een sterk verhoogde kans op het optreden van incongruentiearthrose.

4) Aangezien het functioneren van gewrichtsvlakken onderling een trophische invloed heeft op het bekledend kraakbeen, geeft alleen een behandeling, die spoedig gewrichtsoefeningen mogelijk doet zijn, een optimale genezingskans voor het belangrijke gewrichtskraakbeen.

In het volgend onderdeel van het eerste hoofdstuk werden de verschillende indelingen van de fracturen van het B.S.G. besproken.

Het indelingsschema van Weber wordt gegeven. Hiervan zijn voor dit onderzoek vooral de A-, B- en C-fracturen uit de eerste groep fracturen uit de indeling van Weber van belang.

Deze eerste groep betreft de z.g. luxatiefracturen. Met deze indeling, zo stelt Weber is de prognose van het letsel bepaald, aangezien bij de A fracturen nooit, bij de B-fracturen vaak, en bij de C-fracturen altijd de voorste syndesmose is verscheurd.

In de daarna volgende onderdelen van het eerste hoofdstuk worden de operatie-indicaties, de operatietechnieken en de naonderzoekmethoden die door Weber worden geadviseerd en beschreven, uitgebreid belicht.

In hoofdstuk II worden de gegevens die werden verzameld bij het naonderzoek van de primair volgens Weber geopereerde groep (PO) volgens de door Weber aangegeven toetsingscriteria vergeleken met de gegevens van de patiënten uit de niet geopereerde groep (NO). Hierbij bleek de PO groep als geheel een beter resultaat te scoren dan de NO groep. Bij een onderverdeling in A-, B- en C-fracturen bleek er tussen de geopereerde patiënten met een A-fractuur en de vergelijkbare niet geopereerde patiënten geen meetbaar resultaatverschil te bestaan. Bij de B- en C-fracturen bleken de geopereerde patiënten echter steeds beter. Het leek nu voor de hand te liggen om de operatieve behandeling tot de betere behandeling te verklaren. Het bleek echter dat vergelijking van de geopereerde patiënten uit het onderzoek met de resultaten van de door Weber gepubliceerde naonderzoekgegevens, sterk ten nadele van de eigen onderzochte patiënten uitviel. Dit leidde tot een nadere analyse van de mogelijke achtergronden van dit verschil.

Hierbij bleek, dat het tijdsinterval tussen fractuur en moment van naonderzoek bij Weber gemiddeld 63 weken was, terwijl de PO groep na gemiddeld 181 weken werd naonderzocht. Complicaties van een gewrichtsletsel zoals posttraumatische arthrose hebben tijd nodig om waarneembaar te ontstaan. Hoe snel of hoe langzaam deze processen verlopen is een open vraag, dus bij een verschil in tijd is de vergelijk-

baarheid van de naonderzoekgegevens uiterst dubieus. Helemaal onzeker is het welke conclusies uit een dergelijke vergelijking getrokken mogen worden.

De patiënten uit de NO groep werden echter gemiddeld na ruim 300 weken onderzocht, zodat het er naar ging uitzien of er in het geheel geen verschil in de behandelingsresultaten van de drie groepen zou kunnen worden aangetoond, doch dat alle schijnbare verschillen berusten op een langzaam in de tijd voortschrijdend arthrotisch proces, dat op drie belangrijk van elkaar verschillende tijdstippen was waargenomen.

Nadere analyse van deze foutenbron bij het beoordelen van klinische naonderzoekgegevens leidt tot het volgende hoofdstuk (III) waarin deze en nog een vijftal andere belangrijke foutenbronnen worden besproken, en wel:

- gebruik van ongelijke klassificaties
- verkeerd en ongetoetst gebruik van hypothesen
- onduidelijk gedefiniëerde beoordelingsnormen of toetsingscriteria
- onjuiste of onduidelijke kwantificaties
- invloed van selectie

Bewerking van klinische naonderzoekgegevens sluit vrijwel altijd het gebruik van statistische beslissingsmethoden in. De principes hiervan, met zijn mogelijkheden en beperkingen worden dan ook in dit hoofdstuk besproken.

In hoofdstuk IV worden de bij het onderzoek verzamelde gegevens en de wijze van kwantificering besproken.

In hoofdstuk V komen de vergelijkbaarheid, de representativiteit en de bij de verdere bewerking gebruikte toetsingscriteria aan de orde. Het blijkt, dat met in acht nemen van een aantal interpretatierestricties de PO en NO groep toch redelijk vergelijkbaar zijn ondanks het grote verschil in naonderzoekperiode.

Wat de toetsingscriteria betreft wordt erop gewezen dat het hierbij van belang is de objectieve waarnemingsgegevens gescheiden te houden van de gegevens betreffende het subjectieve welbevinden van de patiënt, aangezien de relaties tussen deze twee groepen gegevens meestal onbekend zijn en het tenslotte om het subjectieve welbevinden van de patiënt en de prognose daarvan gaat als het de bedoeling is het resultaat van een behandeling vast te stellen.

Dit brengt de verplichting mee om tenminste met twee toetsings criteria te werken. Hiervan vertegenwoordigt het ene het subjectief welbevinden en het andere het totaal aan objectieve waarnemingsgegevens.

Statistische beslissingsmethoden stellen in staat generaliserende uitspraken te doen aan de hand van bestudering van steekproeven. Voor welke patiëntengroepen deze generalisaties geldig kunnen zijn is de vraag, die aan de orde wordt gesteld bij de bespreking van de representativiteit van het onderzoek. Er wordt op gewezen dat de generalisatiegeldigheid verstoord wordt door het overlijden van de oudere patiënten groep voordat ze aan naonderzoek zijn onderworpen. Ook zijn er patiënten die wel

in de steekproef thuishoren, doch door verhuizing of andere redenen niet voor een naonderzoek bereikt kunnen worden. Het effect op de beoordeelbaarheid hiervan alsmede een methode om dit naonderzoekverlies op zijn betekenis te schatten wordt besproken.

Hoofdstuk VI is geheel gewijd aan de betekenis van het tijdsverloop tussen oorspronkelijke afwijking en het moment van naonderzoek. Er wordt gezocht naar samenhang tussen de waarde van de onderzoekvariabelen en de grootte van dit tijdinterval (naonderzoekperiode). Bij een aantal variabelen wordt een duidelijke relatie met deze naonderzoekperiode gevonden. Dit toont aan dat in principe slechts op twee wijzen nauwkeurige conclusies te trekken zijn uit de vergelijking van twee groepen klinische gegevens:

- De twee groepen dienen transversaal in de tijd te zijn onderzocht. D.w.z. de naonderzoekperiode van de patiënten uit de twee groepen moet praktisch identiek zijn, of
- Dezelfde patiënten, worden meerdere malen onderzocht. D.w.z. een longitudinaal onderzoek in de tijd, waarbij groepen patiëntengegevens worden vergeleken, die alleen in naonderzoek periode verschillen. Een dergelijk onderzoek is nodig indien juist het nagaan van de veranderingen in de tijd het doel van het onderzoek is. Gegevens uit dergelijke onderzoeken zijn onmisbaar voor het nauwkeurig leren stellen van prognoses.

Voor dit onderzoek heeft de tijdafhankelijkheid van een aantal variabelen de consequenties dat bij het interpreteren van de bewerkingsuitkomsten een aantal restricties in acht moeten worden genomen.

Hoofdstuk VII begint dan ook, na het nog eens formuleren van het onderzoek doel, met het opstellen van de in acht te nemen interpretatierestricties.

De rest van dit hoofdstuk handelt over de bewerking van de onderzoekgegevens en de conclusies die daaruit moeten worden getrokken. Bij de bespreking van de somscore van alle objectieve waarnemingsgegevens was reeds naar voren gekomen dat de niet röntgenologische gegevens, die in hoofdzaak van het beoordelen van de weke delen toestand rond het enkelgewricht afkomstig zijn, zich anders leken te gedragen dan de gegevens, die afkomstig zijn van de toestand van de skelet elementen van het bovenste spronggewricht, voor zover die op de röntgenfoto werd beoordeeld. Om deze reden werd de totale objectieve somscore verdeeld in twee scores. Eén ervan (178) was de somscore van alle fysisch diagnostische gegevens, die in hoofdzaak de toestand van de weke delen vertegenwoordigt. De andere (179) vertegenwoordigt de toestand van het benige skelet van het B.S.G. beoordeeld aan de hand van de röntgenfoto.

Als derde toetsingskriterium werd een subjectieve somscore gebruikt (97), die het subjectieve welbevinden van de patiënt weergeeft.

Het gedrag van de toetsingskriteria werd nagegaan, waarbij bleek, dat er een

nauwere samenhang bestond tussen de weke delen verhoudingen (178) en het subjectief welbevinden (97), dan tussen de röntgenologisch waarneembare afwijkingen (179) en 97. Bovendien bleek er een afhankelijkheid te bestaan tussen 178 en de grootte van het tijdsinterval tussen fractuur en naonderzoek (87).

Het bleek nu dat:

- Indien de totale PO en NO groepen met elkaar werden vergeleken dan zijn zowel score 97 en 179 bij PO gunstiger, doch 178 juist ongunstiger.
- Indien subgroepen worden gevormd van A-, B- en C-groepen, dan blijken de verhoudingen bij de B-fracturen identiek te zijn aan die bij de totale groepen. D.w.z. alle toetsingskriteria verschillen significant tussen de PO en NO groep, doch alleen de weke delen situatie is bij de geopereerde groep significant ongunstiger.

Bij de A- fracturen is geen enkel verschil aantoonbaar, terwijl bij de C-fracturen alleen de röntgenologische verhoudingen bij de geopereerde groep veel gunstiger zijn, terwijl subjectief geen verschil blijkt te bestaan.

Om een uitspraak te kunnen doen over het al of niet beter zijn van een bepaald resultaat worden de volgende voorwaarden gesteld:

- Het 'betere' resultaat dient in gunstige zin significant te verschillen ten aanzien van het subjectieve welbevinden (97) van de patiënten van de vergeleken groep.
- In verband met de sterke samenhang tussen de weke delen verhoudingen en het subjectief welbevinden van de patiënten alsmede door de aantoonbare relatie tussen de tijd en score 178, waaruit moet worden geconcludeerd, dat de weke delen verhoudingen geleidelijk steeds ongunstiger worden, dient het 'betere' resultaat na dezelfde naonderzoekperiode ten aanzien van de weke delen verhoudingen minstens evengoed te zijn als dat van de vergeleken patiënten.
- Ten aanzien van score 179 worden aan het 'betere' resultaat dezelfde voorwaarden gesteld als aan score 178. Ten opzichte van de anatomische röntgenologisch vaststelbare afwijkingen kon geen duidelijke samenhang met de tijd (naonderzoekperiode) worden vastgesteld. Bovendien bleek de samenhang van deze gegevens (179), met het subjectief welbevinden (97) van de onderzochte patiënten minder sterk te zijn dan tussen 178 en 97. Voorlopig wordt toch aangenomen, dat een ongunstige score 179 de prognose van het subjectief welbevinden ongunstig beïnvloedt.

Aan de hand van deze uitspraakvoorwaarden is op grond van de bewerkingsresultaten van de fracturen ingedeeld naar Weber geen uitspraak mogelijk.

Een nauwkeuriger analyse van het gedrag van de toetsingskriteria bracht het volgende aan het licht:

- De indeling van de enkelfracturen volgens Weber geeft onvoldoende informatie over de prognose van het letsel.

- Dit wordt veroorzaakt door het feit dat deze voor een groot deel wordt bepaald door de toestand van de weke delen rond het gewricht, zoals die bij het naonderzoek wordt gevonden.
- Deze uiteindelijke toestand van de weke delen blijkt voor het grootste deel af te hangen van de primaire beschadiging zoals die bij het oorspronkelijke letsel is ontstaan.
- De primaire beschadiging van de weke delen, voorzover hij aanleiding geeft tot late ongunstige verhoudingen, is met grote nauwkeurigheid af te lezen aan de primaire dislocatie zoals die op de ongevalsfoto te vinden is.
- Daarom dient de eerste klassificatie van de luxatiefracturen van het B.S.G. met het oog op het bepalen van de prognose van het letsel, een klassificatie te zijn naar primaire dislocatie.
- Pas daarna komt het aan de orde om te trachten binnen gelijke dislocatiegraden naar subindelingen te zoeken, die een eventueel indicatiegebied voor operatieve of niet operatieve behandeling nauwkeuriger aangeven.  
Het is niet onwaarschijnlijk dat dergelijke subindelingen volgens de Weberklassificatie uitgevoerd, hun waarde nog zullen bewijzen.
- De toestand van de weke delen wordt door een operatieve behandeling hoogstens in ongunstige zin beïnvloed.
- De toestand van het röntgenologisch beoordeelbare gewrichtsskelet is door operatieve behandeling zeer sterk in gunstige zin te beïnvloeden.
- Door deze situatie zit de invloed van de weke delen verhoudingen als het ware op de wip van de operatieindicaties:  
Indien de weke delen verhoudingen primair niet al te ongunstig zijn dan kan het operatief herstellen van de benige verhoudingen en het hechten van de ligamenteuze verbinden nauwkeuriger dan door middel van conservatieve behandeling mogelijk is, voor de patiënt subjectieve winst geven. Doch, indien de primaire weke delen verhoudingen zeer ongunstig zijn, dan heeft operatief herstel geen zin. Immers conservatieve repositie kan de benige relatie zover herstellen, dat overblijvende restdislocaties ten aanzien van het uiteindelijk subjectieve welbevinden een ondergeschikte rol spelen. Dit subjectieve welbevinden wordt dan in sterk overwegende mate bepaald door de weke delen, waarvan de toestand door operatie hoogstens ten nadele verandert.
- Het is opvallend, dat ook bij zeer geringe dislocatie van de operatieve behandeling geen winst mag worden verwacht.  
Het lijkt erop of Weber wat al te sterk mechanisch en te weinig biologisch heeft geredeneerd, immers het volstrekt niet accepteren van incongruenties van gewrichtsvlakken impliceert, dat hij geen enkele spontane compensatie- en genezingscapaciteit in dit opzicht aanneemt van belangrijke structuren als gewrichten.  
Het lijkt waarschijnlijker, dat er wel degelijk een spontane genezingscapaciteit

ook in dit opzicht bestaat en dientengevolge het alleen zinvol kan zijn van buiten af in te grijpen als de dislocaties deze eigen genezingscapaciteit overschrijden.

- Ten slotte bleek uit de onderzoekgegevens, dat alleen bij dislocatiegraad twee de operatieve behandeling volgens Weber voordelen bood ten opzichte van de niet operatieve behandelingstechniek.

Indien de dislocatiegraad groter was dan twee, dan werd bij operatieve behandeling of het röntgenaspect bij de geopereerde patiënten aanzienlijk beter zonder dat het subjectief welbevinden van deze patiënten verschilt van dat van de niet geopereerde patiënten, of ook bij de geopereerde patiënten zijn zoveel secundaire degeneratieve afwijkingen op de röntgenfoto waar te nemen dat helemaal geen enkel verschil tussen geopereerde en niet geopereerde patiënten is vast te stellen.



## SUMMARY

In 1966, Weber advised open reduction and fixation, followed by early exercise, for nearly all dislocation-fractures of the ankle. In this book, a group of 97 patients treated in this way, is compared with a group of about 250 patients treated with closed reduction and immobilization. The first group was examined between two and five years after treatment, the second group between four and eight years. Weber's advice is grounded on a number of anatomical and functional observations that are discussed in chapter I.

The major points are:

1. The talar movement with regard to the tibia is not the simple dorsi-, plantar flexion around a transverse axis, but a complicated movement with a rotation axis in three different planes.
2. Therefore the fibula must move along with the rotating and tumbling lateral side of the astragalus, in order to sustain contact between the articular surfaces. This means that for an optimal function of the ankle joint independent movement of the fibula, with regard to the tibia, is necessary.
3. During certain phases of the gait, the astragalus forces itself against the lateral malleolus.
4. During these phases, the ligaments between the tibia and the fibula prevent a subluxation of the talus.

From these points Weber arrived at the following hypotheses:

1. An absolute stable union between the medial and lateral malleolus, the so called ankle-fork, is mandatory to create an optimal chance of cure in fractures of the ankle joint.
2. The ankle-fork stability is seriously threatened by insufficiency of one of the inferior ligaments between the fibula and the tibia. In dislocation-fractures of the ankle joint, especially the anterior inferior tibiofibular ligament is torn.
3. During normal use of the ankle joint, the forces acting upon the articular surfaces are very great, so that any incongruity increases the chance of later arthrosis.
4. The joint function has a trophical influence on the cartilage of the articular

surfaces. Therefore a speedy recovery of this important cartilage is only possible with a form of treatment allowing early exercise.

In the following part of chapter I the classifications of fractures of the ankle joint are discussed.

Especially Weber's classification of dislocation-fractures of the ankle, is important in determining the prognosis of the ankle injury, because, as Weber states, in A-fractures the anterior inferior tibiofibular ligament is never torn, in B-fractures sometimes, and in C-fractures always.

The last part of chapter I discusses the indications for surgery, the surgical techniques, and the follow up examinations as given by Weber. In chapter II the results of the follow up of the group of primarily operated patients, according to Weber's indications (PO), are compared with the results of the follow up of the group of conservatively treated patients (NO). The comparison was done with specific tests, as given by Weber. In this comparison the PO group scored a better result than the NO group. After separating the fractures into Weber's A-, B- and C fractures no difference was found between A fractures of the PO and NO group. With B- and C-fractures, the PO group scored better. It now seemed a logical conclusion to declare the open reduction and fixation the better method of treatment in dislocation-fractures of the ankle.

However, the PO group results turned out to be inferior to Weber's own published results: consequently further analysis of the possible source of these different results, was undertaken.

This revealed that the average time interval between the injury and the follow up examination in Weber's series was 63 weeks, while in the PO group this was 181 weeks. Complications of injuries to joints, as for instance post traumatic arthrosis, need time to reveal themselves. How fast these processes develop is not yet known, and with a different time interval, the comparison of follow up results is hardly possible. It is even more difficult to know what conclusions can be drawn from such a comparison.

The average time interval for the NO group, was more than 300 weeks, so now it seems that there is no difference at all in the results of treatment in the three mentioned groups, and that all the apparent differences are merely phases in a slowly developing arthrosis. Investigation of other possible explanations for misinterpreting clinical follow up results, led to the following five points discussed in chapter III.

- the use of different classifications.
- the use of wrong and non verified hypotheses.
- vaguely defined evaluation criteria.
- wrong or unspecified quantifications.
- the influence of selection.

The processing of clinical follow up results, inevitably necessitates the use of statistical decision methods. The principles of statistics with their possibilities and limitations are also discussed in this chapter.

In chapter IV, the data collected during the follow up and the methods of quantification are discussed.

In chapter V, the comparability, the representivity of the groups, and the evaluation criteria used, are discussed.

It is found that when certain restrictions with regard to interpretation are made, the PO and NO groups are reasonably comparable, in spite of the difference in the follow up time interval.

The importance of continually separating the objective data from the subjective data concerning the well-being of the patient is stressed, because the relations between these two groups of data are usually unknown. It must be born in mind that the subjective sense of well-being of the patients is really one of the most important parameters in trying to evaluate the result of a certain method of treatment. This implies that two sets of evaluation criteria are necessary. One represents the subjective sense of well-being, the other the objective data.

Statistics allow drawing general conclusions from data collected by sampling. To which group of patients these general conclusions refer is mentioned in the discussion about the representativity of the investigation. It is stressed that generalization can be disturbed by the decease of elderly patients before they can be submitted to the follow up, or by other patients, belonging to the sample, who cannot be reached for one reason or another.

The significance of all this, and a method to evaluate the loss to follow up for these reasons, is discussed.

Chapter VI is about the significance of the time interval between the original injury and the follow up. Some follow up data show a definite correlation with this time interval (follow up period).

This implies that exact conclusions from comparing two groups of clinical data can only be drawn when one of the two following conditions is met:

1. The follow up period of the two groups of patients must be identical.
2. The same patients are seen more than once, so that groups of data gathered at different time intervals are compared. This last type of investigation enables a better understanding of the prognosis.

For the present investigation, the time dependence of some follow up data, implies that certain restrictions have to be made when interpreting the results.

Chapter VII describes the purpose of the present investigation, followed by listing the aforementioned interpretation restrictions.

The other part of this chapter discusses the way the results were used, and the conclusions to be drawn from them.

During the discussion about the total score of all the objective data it was noticed that the skeletal data of the ankle joint, as seen on the X-rays, and the data concerning the soft tissues of the ankle joint, seemed to behave in a different manner. For this reason, the total objective data score was divided as follows:

One (178) being the total score of the physical signs concerning the soft tissues, the other (179) representing the condition of the skeletal parts of the ankle joint, as judged from the X-ray.

The third evaluation test was a subjective data score (97) showing the patients' view on their own well-being.

This gave the following results:

- Comparison of the PO and NO groups showed that both score 97 and score 179 are better in the PO group, but that score 178 is better in the NO group.
- When subgroups are formed according to Weber's A-, B-, and C fractures, the same results as for the total groups are reached when comparing the PO and NO groups of the B fractures. This means that all the evaluation criteria differ significantly between the PO and NO groups, generally in favour of the PO group, with the exception of the soft tissue situation, which is significantly worse in the operated group. With regard to the A fractures no differences can be shown, while the C fractures show a better radiological situation in the operated group, with no differences in the subjective data.

To be able to say that a certain result is better or not better, a number of conditions have to be met:

- Because of the discovered correlation between the soft tissue relations and the subjective well-being of the patients, and the proven time relation of score 178, which means that the soft tissue situation slowly deteriorates, the 'better' result, after a certain follow up period, will have to be at least as good, as far as the soft tissues are concerned, as the results of the other patients of the compared group, after the same follow up period.
- The 'better' result must differ significantly and favourably with regard to the subjective well-being (97) of the patients of the compared group.
- With regard to the radiologically visible anatomical disturbances no time relation could be found. Also, the correlation between these data (179) and the subjective well-being (97) of the examined patients was less pronounced than between 178 and 97. Nevertheless it is assumed that an unfavourable score 179 has a negative influence on the prognosis of the subjective well-being. Therefore, the same conditions as in score 178 are applied to a 'better' result in score 179.

With these conditions it is impossible to draw conclusions from the results of treatment follow up of ankle fractures, subdivided according to Weber.

A deeper analysis of the evaluation criteria showed the following:

- Weber's classification of ankle fractures does not give enough information about the prognosis of the injury.
- The reason for this is that the prognosis is largely determined by the eventual soft tissue situation as found in the follow up.
- The eventual soft tissue situation is primarily dependent on the original damage at the time of injury.
- The original damage of the soft tissues, responsible for the later unfavourable situation, can be accurately read from the primary displacement, as seen on the X-ray at the time of injury.
- Therefore, first of all, dislocation-fractures of the ankle joint, should be classified according to the primary displacement, so that the prognosis of the injury can be determined using this classification.
- Only then a method to subdivide groups with identical dislocations may be searched for, with the intention to provide a more exact indication for open or closed treatment.

It is not unlikely that such subdivisions according to Weber's classification could prove to be useful.

- Surgical treatment has a negative influence on the soft tissue damage.
- Surgical treatment has a positive influence on the appearance of the ankle joint on the X-ray.
- Because of this situation, it is the soft tissue damage that really determines whether to operate or not; c.q. if the soft tissue damage is not too extensive, the exact reconstruction of the ligaments and bony parts of the ankle joint can of course be better achieved with an operation than with closed reduction, and therefore be to the advantage of the patient. If on the other hand the primary soft tissue damage is very extensive, operation only gives additional trauma, while the less perfect result of closed reduction of the bony ankle joint, plays a minor part in the eventual end result. In other words, the soft tissue damage then determines the end result, which the extra trauma from operative treatment could only worsen.
- Even with very slight dislocation no advantage can be expected from open treatment.

It seems that Weber's reasoning was mainly mechanical instead of biological. His absolutely not accepting incongruity of articular surfaces implies that he considers such important structures as the joint, incapable of compensation and regeneration. It is more likely that a regeneration capacity does exist, and that interference by operation only makes sense, when the dislocation goes beyond this regeneration capacity.

- Finally it was shown that only with a dislocation of grade two, Weber's operative treatment has advantages over the non operative treatment.

When the degree of dislocation is higher, either the X-ray appearance is better than in the closed reduction group, without a difference in the sense of well-being, (meaning that the X-ray was treated and not the injury) or the X-ray shows secondary degenerative signs in both groups, so that there is no way to determine any difference between operated and non operated patients.

## LITERATUUR

- Allredge, R. H.: Diastasis of the distal tibiofibular joint and associated lesions.  
J. of Am. Ass. 115, 2136, 1940.
- Anderson, J., Lecocq, J.: Operative treatment of the tibiofibular collateral ligament of the ankle.  
J. B. J. S. 36A, 285, 1954.
- Ashhurst, A. P. C., Bromer, R.: Classification and mechanism of fractures of the leg bones involving the ankle.  
Arch. Surg. 4, 51, 1922.
- Barnett, C. H., Napier, J. R.: The axis of rotation at the ankle joint in man. Its influence upon the form of the talus and the mobility of the fibula.  
J. Anat. 86, 1, 1952.
- Bedogni, C., Bergami, P. L.: Le fracture trimalleolari.  
Arch. Orthop. LXXV, 105, 1962.
- Bellenger, M.: L'ostéoporose posttraumatique. Aspect clinique et méthodes thérapeutique.  
Acta Chir. Belg. et Acta Orthop. Belg. 16, 404, 1950.
- Biström, O.: Conservative treatment of severe ankle fractures.  
Acta Chir. Scand. suppl. 168, 1952.
- Blinov, B. V.: Knöchelbrüche in Verbindung mit der Fraktur des Vorderen oder hinteren Tibiarandes.  
Orthop. Traumat. Prothes. 24, 16, 1963.
- Bogdanov, F. R., Yaralov-Yarlyantz, V. A.: Moderne Methoden der Behandlung der Fussknochen-Frakturen.  
Orthop. Traum. Prothes. 24, 3, 1963.
- Böhler, L.: Die Verrenkungsbrüche im Sprunggelenk.  
Wien. Klin. Wschr. 40, 104, 1927.
- Bonnier, P.: Les fractures du pilon tibial.  
Thèse, Lyon, 1961.
- Bonnin, J. G.: The hypermobile ankle.  
Proc. Roy. Soc. Med. 37, 282, 1948.
- Borchardt, M.: Frakturen der Malleolen.  
Handbuch der praktischen Chirurgie, Bd. 6, 627, Enke, Stuttgart, 1914.
- Brady, Th., Arnold, A.: Aspiration injection treatment for varus sprain of the ankle.  
J. B. J. S. 54A II, 1257, 1972.
- Braune, W., Fischer, O.: Der Gang des Menschen.  
1. Theil. Abb. d. Math.-Phys. Classe d. kgl. Sächs. Ges. d. Wissenschaften, Bd. 21, 151, 1895.
- Braunstein, P. W., Wade, P. A.: Treatment of unstable fractures of the ankle.  
Ann. Surg. 149, 217, 1959.
- Brodie, I. A. O. D., Denham, R. A.: The treatment of unstable ankle fractures.  
56B, 256, 1974.

- Burwell, H. N., Charnley, A. D.: The treatment of displaced fractures at the ankle by rigid internal fixation and early joint movement.  
J. B. J. S. 47B, 634, 1965.
- Callaghan, J. E., Percy, E. C., Ross O'Hill.: Ankle arthrogram.  
J. Canad. A Radiologists 21, 74, 1970.
- Cedell, C. A., Wiberg, G.: Treatment of eversion-supination-fracture of ankle (2d degree).  
Acta Chir. Scand. 124, 41, 1962.
- Champonnière, J. L.: Traitement des fractures par le massage et la mobilisation.  
Paris 1893 (zit.n. Steinmann, Matti).
- Close, J. R. : Some applications of the functional anatomy of the ankle joint.  
J. B. J. S. 38A 761, 1956.
- Colton, C. L.: Fracture-diastasis of the inferior tibio-fibular joint.  
J. B. J. S. 50B, 830, 1968.
- Colton, C. L.: The treatment of dupuytren's fracture-dislocation of the ankle.  
J. B. J. S. 53B, 63, 1971.
- Coonrad, R., Bugg, J. E.: Trapping of the posterior tibial tendon and interposition of soft tissue in severe fractures about the ankle joint.  
J. B. J. S. 36A, 744, 1954.
- Coonrad, R. W.: Fracture-dislocations of the ankle joint with impaction injury of the lateral weight-bearing-surface of the tibia.  
J. B. J. S. 52A, 1337, 1970.
- Costigan, R. C.: Treatment of true widening of ankle mortise.  
Can. med. an. J. 69. 310, 1953.
- Crock & Crock.: The blood supply of the lower limb in man.  
E & S Livingstone Ltd., Edinburgh & London, 1967.
- Danis, R.: Théorie et pratique de l'ostéosynthèse.  
Paris Masson, 1949.
- Decoulx, P., Razemon, J. P., Rouselle, Y.: Fractures du pilon tibial.  
Rev. Chir. orth. 5, 563, 1961.
- Denham, R. A.: Internal fixation for unstable ankle fractures.  
J. B. J. S. 46B, 206, 1964.
- Desenfans, G.: Réflexions au sujet du traitement du diastasis tibiopéronier à la suite d'une dislocation ouverte du tarse.  
Acta orthop. Belge 25, 279, 1959.
- Destot, E.: Diastasis et fracture des malléoles.  
Rev. Chir. 27, 279, 1907.
- Dinstl, K., H. Spängler.: Über luxations fracturen im Bereich des oberen Sprunggelenkes.  
Arch. Orthop. Unfallchirurgie 55, 317, 1963.
- Dziob, J. M.: Ligamentous injuries about the ankle joint.  
Am. J. Surg. 91, 692, 1956.
- Earle.: Simple, succeeded by compound dislocation forward of the inferior extremity of the tibia.  
Lancet 2, 348, 1828/29.
- Eberhard, H. D., Inman, V. T.: et al. Fundamental studies of human locomotion and other informations relating to design of artificial limbs.  
University of California, Berkely, 1947.
- Ferré, G. A., Platt, J. L.: Treatment of lateral ligamentous injuries of ankle.  
Virginia M. Month. 95, 617, 1968.
- Fordyce, A. J. W., Horn, C. V.: Arthrography in recent injuries of the ligaments of the ankle.  
J. B. J. S. 54B, 116, 1972.
- Fowler, P. J., MacDonald, A. G.: Ligamentous osseous injuries of the ankle. A pathomechanical study.  
J.B.J.S. 56A II, 1303, 1974.



- Gay, R., Evrard, J.: Les fractures récentes du pilon tibial chez l'adulte. Rapport de la réunion annuelle de la Société Française d'Orthopédie et de Traumatologie.  
Rev. Chir. Orth. 4, 397, 1963.
- Gillespie, H. S., Boucher, P.: Watson-Jones repair of lateral instability of the ankle.  
J. B. J. S. 53A, 920, 1971.
- Golterman, A. F.: Letsels van het enkelgewricht.  
Dissertatie, 1965.
- Gordon, R. B.: Arthrography of the ankle joint. Experience in one hundred seven studies.  
J. B. J. S. 52A II, 1623, 1970.
- Grath, G. B.: Widening of the ankle mortise.  
Acta Chir. Scand. suppl. 263.
- Gschwend, N.: Die fibularen Bandläsionen. Eine häufig verkannte Folge der Fussverstauchungen.  
Praxis 47, 809, 1958.
- Hackethal, K. H.: Das Südeck-Syndrom.  
Hüthig, Heidelberg, 1958.
- Hansen, S.: Om ankelbrud.  
Madsen, København, 1919.
- Iselin, M., Vellis, H.: La primauté du péroné dans les fractures du cou-de-pied.  
Mém. Acad. Chir. 399, 1961.
- Jansen, W. B. J.: Osteosynthese van fracturen van het bovenste spronggewricht met functionele nabehandeling.  
Dissertatie 1971.
- Johnson, E. W., Boseker, E. H.: Arthrodesis of the ankle.  
Arch. Surg. 97, 766, 1968.
- Jones, W. C., Neal, E. G.: Survey of fractures of medial malleolus.  
Southern Med. J. 55, 1054, 1962.
- Joy, G. Patzakis, M. J., Harvey, J. P.: Precise evaluation of the reduction of severe ankle fractures.  
J. B. J. S. 56A I, 979, 1974.
- Joyce, J. J., Harty, M.: Ankle approaches.  
Orth. Approaches. Sec. I. Lower extremity. Williams and Wilkins, Baltimore, 1961.
- Julius, H. W.: Het recht op gezondheid.  
Tijdschrift v. Soc. Geneesk. 32, 340, 1954.
- Keet, J. G. M.: De late gevolgen van enkelfracturen.  
Dissertatie, 1970.
- Kleiger, B.: The mechanism of ankle injuries.  
J. B. J. S. 38A, 59, 1956.
- Kleiger, B.: The treatment of oblique fractures of the fibula.  
J. B. J. S. 43A, 969, 1961.
- Klossner, O.: Late Results of operative and non-operative treatment of severe ankle fractures.  
Acta Chir. Scand. suppl. 293, 1962.
- Koch, W.: Zur operativen Behandlung von Malleolar-luxations-Frakturen.  
Monatschr. für Unfallheilk. 71, 347, 1968.
- Kristensen, T. B.: Treatment of malleolar fractures, according to Lauge Hansen's method.  
Acta Chir. Scand. 97, 362, 1949.
- Kristensen, T. B.: Fractures of the ankle VI. Follow-up studies.  
Arch. Surg. 73, 112, 1956.
- Küntschner, G.: Stabile Osteosynthese gelenknaher Brüche.  
Zbl. Chir. 82, 1641, 1957.
- Laskin, R. S.: Steinmann-Pin fixation in the treatment of unstable fractures of the ankle.  
J.B.J.S. 56A I, 549, 1974.

- Lauge Hansen, N.: Fractures of the ankle I. Analysis-history survey as the bases of new experimental, roentgenologic and clinical investigations.  
Arch. Surg. 56, 259, 1948.
- Lauge Hansen, N.: 'Ligamentous' ankle fractures. Diagnoses and treatment.  
Acta Chir. Scand. 97, 544, 1949.
- Lauge Hansen, N.: Fractures of the ankle II. Combined experimental surgical and experimental roentgenologic investigations.  
Arch. Surg. 60, 957, 1950.
- Lauge Hansen, N.: Fractures of the ankle IV. Clinical use of genetic roentgen diagnosis and genetic reduction.  
Arch. of Surg. 64, 488, 1952.
- Lauge Hansen, N.: Fractures of the ankle III. Genetic roentgenologic diagnosis of fractures of the ankle.  
Am. J. Roentgenol. 71, 456, 1954.
- Lauge Hansen, N.: Knöchelbrüche und Bandverletzungen des Sprunggelenkes.  
Hefte zur Unfallheilkunde, no. 92, Springer Verlag, Berlin, 1967.
- Magnussen, R.: On the late results in non-operated cases of malleolar fractures. A clinical roentgenological statistical study. Fractures by external rotation.  
Acta Chir. Scand. 9, suppl. 84, 1944.
- Mainland, D.: The use and misuse of statistics in medical publications.  
Clin. pharmac. and Therapeutics 1, 411, 1960.
- Martinez, C.: et al. Résultats du traitement sanglant des fractures malléolaires.  
Rev. Chir. Orthop. 56, 665, 1970.
- Mayer, H.: Die operative Behandlung der Luxationsfracturen des oberen Sprunggelenkes bei gleichzeitiger Sprengung des Ligamentum interosseum zwischen Tibia und Fibula.  
Chirurg 27, 509, 1956.
- Mehrez, M. E. L., Geneidy, S.: Arthrography of the ankle.  
J. B. J. S. 52B, 308, 1970.
- Miller, A. J.: Posterior malleolar fractures.  
J. B. J. S. 56B, 508, 1974.
- Monk, C. J. E.: Injuries to the tibio-fibular ligaments.  
J. B. J. S. 51B, 330, 1969.
- Nava, B. E.: Traumatic dislocation of the tibialis posterior tendon at the ankle.  
J. B. J. S. 50B, 150, 1968.
- Olerud, S.: Subluxation of the ankle without fracture of the fibula.  
J. B. J. S. 53A, 594, 1971.
- Philips, R. S., Balmer, G. A., Monk, C. J. E.: External rotation fractures of the fibular malleolus.  
Brit. J. Surg. 56, 801, 1969.
- Plaue, R., Hinz, P.: Radiologic diagnosis of ankle joint injuries involving the ligaments.  
Fortschr. Geb. Röntgenstrahlen 112, 769, 1970.
- Pott, P.: The surgical works.  
London Hawes, 1775.
- Quenu, E.: Du diastasis de l'articulation tibio-péronière inférieure.  
Rev. Chir. Paris 36, 897, 1907.
- Quigley, T. B.: Fractures and ligament injuries of the ankle.  
Am. J. Surg. 98, 477, 1959.
- Reimers, C.: Die Brüche des Fussnahen Unterschenkel abschnittes Lagenbeck.  
Arch. klin. Chir. 276, 260, 1953.
- Rieunau, G., Gay, R.: Les pseudarthrosis de la malléole interne.  
Chir. Orthop. 42, 75, 1956.
- Rubin, G., Witten, M.: The talar-tilt ankle and the fibular collateral ligaments. A method for the determination of talar tilt.  
J. B. J. S. 42A, 311, 1960.

- Rümke.: Psychiatrie, deel 1, 2e druk, p.55.  
Scheltema & Holkema N.V., 1957.
- Ruth, C. J.: The surgical treatment of injuries of the fibular collateral ligaments of the ankle.  
J. B. J. S. 43A 1, 229, 1961.
- Sneppen, O.: Long-term course in 119 cases of pseudarthrosis of the medial malleolus.  
Acta orth. Scand. 40, 807, 1970.
- Solonen, K. A., L. Lautamus.: Operative treatment of ankle fractures. 1968.
- Souquet, R.: Fractures du cou-de-pied.  
Editions Doin, 1965.
- Staples, O. S.: Rupture of the fibular collateral ligaments of the ankle; result study of immediate surgical repair.  
J. B. J. S. 56A 11, 1539, 1974.
- Stören, G.: Conservative treatment of ankle fractures.  
Acta Chir. Scand. 128, 45, 1964.
- Sullivan, R. D., Watkins, E., Oberfield, R. A., Khazel, A. M.: Current status of protracted arterial infusion cancer chemotherapy for the treatment of solid tumors.  
S.Cl. of N. Amer. 47, 769, 1967.
- Thorndike, A.: Frequence and nature of sport injuries.  
Am. J. Surg. 98, 316, 1959.
- Trojan, E.: Die Behandlung der Knöchelbrüche mit Abscherung eines grossen hinteren Schienbeinkeiles.  
Ztschr. Orthop. 84, 636, 1954.
- Trojan, E., Jahna, H.: Behandlung und Nachuntersuchungsergebnisse von Knöchelbrüchen mit vorderem Schienbeinkeil.  
Hefte Unfallheilk. 92, 79, 1967.
- Vasli, S.: Surgical treatment of malleolar fractures.  
Acta Chir. Scand. suppl. 226, 1951.
- Viernstein, K., Jantzen, P. M.: Die Verletzungen im bereich des oberen Sprunggelenkes.  
Zeitschr. f. Orthop. 88, 87, 1957.
- Vogel, P. L. de.: Oefenstabiele osteosynthese bij enkelfracturen.  
N.T.v.G. 112, 721, 1968.
- Vogel, P. L. de.: Preliminary results of exercise-stable osteosynthesis of fractures of the ankle.  
Arch. Chir. Neerl. 20, 273, 1968.
- Vogel, P. L. de.: Enige functioneel-anatomische aspecten van het bovenste spronggewricht.  
Dissertatie, 1970.
- Watson-Jones, Sir R.: Fractures and joint injuries.  
Fourth Edition E & Livingstone Ltd., 1955.
- Weber, B. G.: Die Verletzungen des oberen Sprunggelenkes.  
Verlag Hans Huber Bern und Stuttgart, 1966.
- Wilson, F. C., Jan, Skilbred, L. A.: Longterm results in the treatment of displaced bimalleolar fractures.  
J. B. J. S., 48A, 1065, 1966.

## II

- Anderson, T. W.: An introduction to multivariate statistical analysis.  
John Wiley & Sons Inc., New York, 1958.
- Bradford Hill, A.: Principles of medical statistics.  
The Lancet Ltd., 1937.
- Bronowski, J.: The common sense of science.  
Random House Tnc., 1959.
- Cochran, W. G., Cox, G. M.: Experimental design.  
John Wiley & Sons Inc., New York, 1950.

- Croxton, F. E.: Elementary statistics with applications in medicine and the biological sciences.  
Dover Publications Inc., New York, 1959.
- Cruchet, R.: De la méthode en médecine.  
Presses univ. de France, Paris, 1951.
- Fisher, R. A.: On the mathematical foundations of theoretical statistics.  
Philosophical Transactions of the Royal Society, London, Series A, Vol. 222, 309, 1922.
- Fisher, R. A.: The design of experiments.  
Oliver and Boyd, 1935.
- Freudenthal, H.: Waarschijnlijkheid en statistiek.  
Erven F. Bohn N. V., Haarlem, 1966.
- Graybill, F. A.: An introduction to linear statistical models.  
McGraw-Hill book comp. Inc., New York, 1961.
- Jonge, H. de: Inleiding tot de medische statistiek. 2e druk  
Wolters-Noordhoff N.V. Groningen 1963.
- Ledley R. S.: Use of computers in biology and medicine  
McGraw-Hill bookcompany 1965.
- MacKinsey, J. C. C.: Introduction to the theory of games.  
New York, McGraw-Hill, 1952.
- Mainland, D.: Elementary Medical statistics. The principles of quantitative Medicine.  
W. B. Saunders Company, 1952.
- Mainland, D.: The use and misuse of statistics in medical publications.  
Cl. pharmacol. and therap. 1, 411, 1960.
- Martini, P.: Methodenlehre der therap. klin. Forschung.  
Berlin, 1953.
- Mood, A. M., Graybill, F. A.: Introduction to the theory of statistics.  
McGraw-Hill book company, stud. ed., 1963.
- Neiss, F.: Determinanten und Matrizen.  
Springer Verlag, 1967.
- Popper, K. R.: The logic of scientific discovery.  
Oxford univ. press., 1959.
- Popper, K. R.: Objective knowledge. An evolutionary approach.  
Oxf. univ. press, 1974.
- Siegel, S.: Nonparametric Statistics for the behavioral sciences.  
Student Ed., McGraw-Hill Kogakusna Ltd., 1956.
- Rao C. R.: Advanced statistical methods in biometric research.  
John Wiley & Sons. Inc. New York, 1952.
- Reichenbach, H.: Wahrscheinlichkeitslehre.  
Leiden, A. W. Seithof's Uitg. Mij. N.V., 1935.
- Trotter, W.: Observation and experiment and their use in medical sciences.  
Brit. med. J. 26, VI I, 1930.
- Trotter, W.: The collected papers of Winfred Trotter.  
Oxf. univ. press, 1946.
- Weele, L. Th. van der, Sluis, D. M. van der, Meer, Tj. van der: Handleiding WESP.  
Rekencentrum R.U. Groningen, 1974.
- Weiss, L.: Statistical decision theory.  
McGraw-Hill book Company Inc., New York, 1961.
- Wibaut, F.: De methode der geneeskunde.  
Erven F. Bohn N.V., Haarlem, 1962.
- Wilson, E. B. jr.: An introduction to scientific research.  
McGraw-Hill book Company Inc., 1952.

Variabele Nr.	Betekenis
1	1=PO 2=NO
2	O=Vrouw 1=Man
3	Gecompliceerd 0=Neen 1=Ja
4	Leeftijd in jaren
5	Beroep, belasting t.o.v. enkel (rangvolgorde)
6	Vrije tijdsbelasting t.o.v. enkel (rangvolgorde)
7	Jaar ongeval
8	Dagnummer ongeval
9	Links 1=Ja
10	Rechts 1=Ja
11	Fractuur indeling mogelijk $\Delta$ =Ja, 0=Alleen vlg. Weber, 1=L.H., 2=Nee
12	A fractuur
13	B fractuur
14	C fractuur
15	A fractuur
16	B fractuur
17	C fractuur
18	A fractuur
19	B fractuur
20	C fractuur
21	SE 1 fractuur
22	SE 2 fractuur
23	SE 3 fractuur
24	SE 4 fractuur
25	PE 1 fractuur
26	PE 2 fractuur
27	PE 3 fractuur
28	PE 4 fractuur
29	SA 1 fractuur
30	SA 2 fractuur
31	PA 1 fractuur
32	PA 2 fractuur
33	PA 3 fractuur
34	PD 1 fractuur
35	PD 2 fractuur
36	PD 3 fractuur
37	PD 4 fractuur
38	Tertius fragment 1 = Gering 2=tussen $\frac{1}{3}$ en $\frac{1}{5}$ 3 = groter dan $\frac{1}{3}$
39	Talus 1 = Gering 2 = Ernstig
40	Gommunitief mediaal 1 = Gering 2 = Ernstig
41	Gommunitief lateraal 1 = Gering 2 = Ernstig
42	Bandletsel voorste syndesmose 1 = Ja
43	Bandletsel achterste syndesmose 1 = Ja
44	Bandletsel mediaal (rangvolgorde)
45	Bandletsel lateraal (rangvolgorde)
46	Operatie 1 = Ja
47	Volgens Weber 1 = Ja
48	Interval tussen binnenkomst en operatie (uren)
49	Interval tussen ongeval en operatie (uren)
50	Dagnummer operatie
51	Operateur (0 = auteur 1 = ander)
52	Bloedleegte, in kwartieren
53	Moeilijkheid operatie (rangvolgorde)
54	Syndesmoseschroef 1 = Ja
55	Genezing per primam 0 = Ja
56	Secundaire ingreep nodig 1 = Ja
57	Complicaties 1 = Ja
58	Complicatiecode

59	Direct anatomisch resultaat (rangvolgorde)
60	Luxatiestand naar lateraal (in millimeters)
61	Luxatiestand naar mediaal (in millimeters)
62	Luxatiestand naar voren (in millimeters)
63	Luxatiestand naar achteren (in millimeters)
64	Varus talusrol (in graden)
65	Valgus talusrol (in graden)
66	Trapje in gewrichtsvlak t.h.v. tertiusfragment (in millimeters)
67	Primaire dislocatie (rangvolgorde)
68	Enkelvork te breed (in millimeters)
69	Operatieoefenstabiel 0 = Ja
70	Intercurrerend overleden 1 = Ja
71	
t/m	Werkhervatting
74	
75	% werkhervatting
76	na ..... weken.
77	100% werkhervatting na ..... weken
78	lopen zonder klachten 1 = Nee
79	na ..... weken.
80	Loopklachten (rangvolgorde)
81	Eind behandeling na ..... weken
82	Osteosynthese materiaal verwijderd in weken, na osteosynthese
83	Onder lokaal anaesthesie 1 = Nee
84	Onder narcose 1 = Ja
85	Genezing per primam 1 = Nee
86	Secundaire ingreep nodig 1 = Ja
87	Nacontrole in weken
88	Pijn (rangvolgorde)
89	Lopen (rangvolgorde)
90	Activiteit
91	Anatomie (Röntgen beoordeling)
92	Functie B.S.G. (rangvolgorde)
93	Functie OSG (rangvolgorde)
94	Totaal score van Weber
95	Eindbeoordeling van Weber
96	Patiënt subjectief tevreden (rangvolgorde)
97	Subjectieve somscore
98	Paraachillairrelief(rangvolgorde)
99	Symmetrie van var. 98
100	Lengtegewelf (rangvolgorde)
101	Holvoet 1 = Ja
102	Spreadvoet 1 = Ja
103	Superductusstand tenen 1=Ja
104	Symmetrie voeten
105	Oud/ander letsel voet heterolateraal 1 = Ja
106	Oud/ander letsel voet homolateraal 1 = Ja
107	Oud/ander been heterolateraal 1 = Ja
108	Oud/ander letsel been homolateraal 1 = Ja
109	Pathologische vaattekening homolateraal 1 = Ja
110	Pathologische vaattekening heterolateraal 1 = Ja
111	Pigmentaties been homolateraal 1 = Ja
112	Pigmentaties been heterolateraal 1 = Ja
113	Pigmentaties operatielitteken 1 = Ja, matig 2 = ja, sterk
114	Functie VSG, verschil met de gezonde kant in tienden.
115	Cysteuze zwelling over voorste syndesmoose 1 = Ja, 2 = Ja, pijnlijk
116	"Bandgevoel" 1 = ja
117	"Slechtweer klachten" 1 = Ja
118	"Start pijn" 1 = even, 2 = matig, 3 = sterk
119	Weefsel praeachillair (rangvolgorde)

120	Drukpijn mediale band, 1 = Ja
121	Drukpijn laterale band, 1 = Ja
122	Drukpijn voorste syndesmose 1 = Ja
123	Drukpijn achterste syndesmose 1 = Ja
124	Drukpijn mediale malleolus 1 = Ja
125	Drukpijn laterale malleolus 1 = Ja
126	Restdislocatie laterale malleolus (in millimeters)
127	Restdislocatie mediale malleolus (in millimeters)
128	Restdislocatie in mediale tibiahoek (in millimeters)
129	Restdislocatie tertiusfragment, trapje gewrichtsvlak (in mm)
130	" " " , spleet in gewrichtsvlak ( in mm)
131	Restluxaties, naar lateraal ( in mm)
132	Restluxaties, naar mediaal ( in mm)
133	Restluxaties, naar voren ( in mm)
134	Restluxaties, naar achteren ( in mm)
135	Valgus talusrol (in graden)
136	Varus talusrol (in graden)
137	Vork te wijd ( in mm)
138	Sclerose gewrichtsvlak, mediaal (rangvolgorde)
139	Sclerose gewrichtsvlak, lateraal (rangvolgorde)
140	Vernauwing gewrichtsspleet, mediaal (in mm)
141	Vernauwing gewrichtsspleet, lateraal ( in mm)
142	Vernauwing gewrichtsspleet, voor (in mm)
143	Vernauwing gewrichtsspleet, achter (in mm)
144	Vernauwing gewrichtsspleet, tussen malleolus lat. en talus (in mm)
145	Vernauwing gewrichtsspleet, tussen malleolus med. en talus (in mm)
146	Vernauwing gewrichtsspleet, tussen tibia en fibula (in mm)
147	Onregelmatig gewrichtsvlak, mediaal (rangvolgorde)
148	Onregelmatig gewrichtsvlak, lateraal (rangvolgorde)
149	Onregelmatig gewrichtsvlak, voorkant(rangvolgorde)
150	Onregelmatig gewrichtsvlak, achterkant (rangvolgorde)
151	Onregelmatig gewrichtsvlak, tussen malleolus med. en talus (rangvolgorde)
152	Onregelmatig gewrichtsvlak, tussen malleolus lat. en talus (rangvolgorde)
153	Onregelmatig gewrichtsvlak, tussen tibia en fibula (rangvolgorde)
154	Osteophytvorming, voorzijde tibia,
155	Osteophytvorming, achterzijde tibia,
156	Osteophytvorming, malleolus medialis,
157	Osteophytvorming, malleolus lateralis,
158	Botstructuur, tibia 1 = verstoorde botstructuur
159	Botstructuur, tibia krachtlijnen 1 = Ja
160	Botstructuur, malleolus lateralis 1 = verstoorde structuur
161	Botstructuur, malleolus lateralis krachtlijnen 1 = Ja
162	Botstructuur, malleolus medialis 1 = verstoorde structuur
163	Botstructuur, malleolus medialis krachtlijnen 1 = Ja
164	Bandverkalking, mediale band 1=Ja
165	Bandverkalking, laterale band 1 = Ja
166	Bandverkalking, voorzijde gewricht 1 = Ja
167	Bandverkalking, achterzijde gewricht 1 = Ja
168	Verbening ter hoogte van syndesmoseschroef 1 = Ja, 2 = synostose
169	Stabiliteit banden, mediaal (stresshoek in graden)
170	Stabiliteit banden, lateraal (stresshoek in graden)
171	Stabiliteit banden, vork te wijd (in mm)
172	Stabiliteit banden, voor-achterwaarts, (in mm)
173	Verkalking tibio-fibulair
174	Follow-up periode in klassen van 52 weken.
175	Objectieve somscore
178	Somscore van de niet röntgenologische objectieve naonderzoekgegevens.
179	Somscore van de röntgenologische naonderzoekgegevens.